NON-PERTURBATIVE TECHNIQUES

IN GAUGE THEORIES

Thesis

Submitted by

PETER VICTOR DUNSFORD SWIFT

for the degree of

DOCTOR OF PHILOSOPHY

Department of Physics

University of Edinburgh

AUGUST, 1984.

TO MY PARENTS

ACKNOWLEDGEMENTS

I wish to express my thanks to my supervisors, Professors David Wallace and Peter Higgs, for their advice and encouragement.

I should like to thank Christopher Bishop for his stimulating collaboration. I am also grateful to Harry Kogan and Craig Smith for their assistance with the numerical work in Chapters 2 and 3.

For her fast and efficient typing I would like to thank Mrs. Ray Chester.

Financial support from the Science and Engineering Research Council is gratefully acknowledged.

Finally I should like to thank my friends in Edinburgh for an enjoyable three years in the Theoretical Physics group.

ABSTRACT

We consider two techniques used to obtain non-perturbative results in gauge theories: the semi-classical approximation and the lattice discretization of space-time.

Application of the semi-classical technique to gauge theories yields a description in terms of instantons, the finite action solutions of the Euclidean field equations. The expression for the interaction action for an instanton in a weak external field is shown to hold for cases where the external field is not necessarily weak. We extend the idea of populating an external field with a gas of instantons by requiring that these decorating instantons be dressed themselves in a self-consistent way. This gives rise to a new effective coupling, similar to the old one, but possessing slightly different qualitative and quantitative features, in particular the onset of the crossover from weak to strong coupling becomes even sharper.

We calculate the probability distribution function for the topological charge contained in a sphere of finite radius due to a dilute gas of instantons. We work self-consistently within the dilute gas approximation and do not need to introduce an arbitrary cutoff on instanton scale sizes. Monte Carlo simulations reflect qualitatively all the features of our results and new features are predicted which may be tested in future simulations.

The lattice regularization of gauge theories is a powerful technique for calculating non-perturbative effects but there are many theories for which appropriate lattice analogues of the action

are hard to find. We present a class of theories with left-right asymmetry which escape the no-go theorems by using the Higgs mechanism to generate Wilson fermions. We introduce a lattice action for the electroweak SU(2) × U(1) gauge theory which has all the features of the continuum model, including a low energy theory analogous to the Fermi theory.

CONTENTS

CONTENTS (Contd.)

PROLOGUE

One of the most widely used techniques to obtain quantitative results in quantum field theory has traditionally been perturbation theory. However, as long ago as 1952 Dyson conjectured that the perturbation series for quantum electrodynamics, the U(1) gauge theory of electromagnetic interactions, was divergent. Consequently the perturbation series is certainly not guaranteed to describe all the physics of the model. These ideas are not specific to the U(1) gauge theory: field theories in general do not have convergent perturbation series (see Jaffe, 1965 and references therein).

The perturbation series is not, however, entirely useless. When the value of the expansion parameter (usually the coupling coefficient) is small - of the order of $\frac{1}{n}$, say, - then typically up to the first n terms of the series may be used to obtain good quantitative results: after this many terms the series begins to diverge manifestly and the results are not trustworthy. This explains the astounding successes of the technique in quantum electrodynamics. Nevertheless, the perturbation series has two great deficiencies. It cannot be used to describe strong coupling physics (in which, typically, the coupling $g > 1$) at all and so a study of the confinement mechanism in quantum chromodynamics, the SU(3) gauge theory of strong nuclear reactions, is not possible. Furthermore, there are non-perturbative effects such as quantum tunnelling, which modify the vacuum of the theory. These can give non-zero effects even when the value of the coupling is small, but as they are zero to all orders of perturbation theory they do not appear in the perturbation series.

The formulation of quantum field theory in terms of a functional integral followed Feynman's formulation of quantum mechanics in terms of a sum over all appropriate paths (Feynman, 1948). For the field theory the dynamical variables are the fields $\phi_i(x,t)$ functions of the space-time coordinates, and we integrate over the fields ϕ_i :-

$$Z = \int \mathcal{D}\phi_1 \ldots \mathcal{D}\phi_n \, e^{\frac{i}{\hbar} S[\phi_1,\ldots,\phi_n]} \qquad (1)$$

for a theory with n fields. S is the classical action. Now as the action is real, the integrand in (1) is an oscillating exponential and so the functional integral is not well-defined. The usual way to overcome this problem is to formulate the theory in Euclidean space-time (Fradkin, 1959; Nakomo, 1959; Schwinger, 1959) by making a Wick rotation and introducing an imaginary time co-ordinate τ

$$\tau = it .$$

The action picks up a factor i and we have

$$Z = \int \mathcal{D}\phi_1,\ldots, \mathcal{D}\phi_n \, e^{-\frac{1}{\hbar} S_E[\phi_1,\ldots, \phi_n]} \qquad (2)$$

with S_E being the Euclidean action. There is a strong connection between this formulation of Euclidean quantum field theory and statistical mechanics - (2) may be thought of as the partition function for a statistical system with energy functional $S_E[\phi_1,\ldots, \phi_n]$ and with $\hbar$ playing the role of $k_B T$. There is a clear correspondence between quantum uncertainty (we must sum over all trajectories, not just the classical one) and thermal fluctuations (we must go beyond mean field theory and construct ensemble averages). From now on we shall adopt units so that $\hbar = 1$.

The Euclidean space-time formulation has an extra significance when we consider quantum tunnelling phenomena. In classically allowed regions no tunnelling takes place and the functional integral (1) is dominated by configurations which are solutions of the classical field equations. However, in the regions which are forbidden classically there are tunnelling solutions which are solutions of the field equations in imaginary time (McLaughlin, 1972). In these regions the functional integral (2) is dominated by these (finite action) Euclidean solutions, referred to as instantons. The obvious reason why the functional integral is dominated by such finite action solutions is that the suppression factor in the integrand from e^{-S_E} is non-zero; however this is not sufficient as such configurations form a set of measure zero in the functional integral. Their importance comes from making semi-classical expansions to approximate the integral. Only expansions made about points with finite action will give non-zero results (Coleman, 1977).

In Part I of this thesis we examine some features of instantons in non-Abelian gauge theories. Chapter 1 contains a discussion and analysis of the vacuum structure and the dilute gas approximation, widely used in instanton calculations. Application of one loop renormalization yields a density function for instantons which grows with increasing scale size. This infra-red divergence causes the eventual breakdown of the dilute gas approximation and is a reflection of the onset of strong coupling physics. The standard way to cope with this problem has been to impose an arbitrary maximum on the scale size of the instantons: however we present in this thesis two calculations in which no such arbitrary cut-off is needed. In Chapter 2 we discuss the renormalization of an external field due to the presence of instantons within that field. There is no need for an arbitrary cut-off as we only include instantons up to the scale size of the external field. The effect of placing instantons within instantons is considered - this produces a self-consistent dressing of the external field. We find that this causes a further anti-screening effect and so the crossover from weak to strong coupling becomes even sharper. In Chapter 3 we calculate the probability distribution function for the topological charge contained in a sphere of finite radius and compare our results with computer simulations of the gauge field. Again, no arbitrary cut-off is needed as only instantons of a certain rigidly controlled range of scale sizes can contribute to the quantity we calculate.

The other main non-perturbative technique currently being studied is the lattice gauge theory. The functional integral is attacked by discretizing the space-time, yielding a countable product of integrals

$$\int \mathcal{D}\phi \quad \rightarrow \quad \prod_n \int d\phi(n)$$

where n labels the sites of the lattice. The introduction of the lattice causes a fundamental violation of Lorentz invariance. We must formulate the theory in Euclidean space-time, as already described, so that we are left with a discrete rotation and translation group after the introduction of the lattice. Then as we take the continuum limit we must hope that the Euclidean translation and rotation group will be restored, giving Lorentz invariance after rotating back to Minkowski space.

If this lattice is placed in a finite space-time box we will have a finite number of degrees of freedom, thus enabling a computer simulation of the model. At first we might think of enumerating all possible field configurations and averaging over them with the appropriate Boltzmann weight: however such an approach is not possible in practice, due to the enormous number of configurations. However, an alternative exists: the Monte Carlo technique. We only need average (with no weight factor) over a relatively small number of "important" configurations, which correspond to thermodynamic equilibrium (see Binder, 1979). Unfortunately there is a difficulty with this approach. There is no guarantee that the "important area" of field configurations is connected in the gauge field space, and most algorithms for generating new configurations take exponentially many steps to cross from one disconnected region to another. Consequently, the system may tend to lie in metastable states rather than true equilibrium states, jumping from one to another after a very large number of steps in each.

The lattice also provides a regularization of the model. This is a great advantage as the theory is then well-defined from the beginning even though we are dealing with bare quantities. Another fundamental advantage of the lattice is that it permits us to examine strong coupling physics, inaccessible from either perturbation theory or the semi-classical approximation. However, there are disadvantages associated with lattice gauge models and, of these, some of the most serious involve fundamental difficulties in the satisfactory formulation of suitable lattice actions.

We turn our attention to this problem in Part II of this thesis. Chapter 4 introduces lattice gauge theories and outlines the notorious fermion species doubling problem which lies at the heart of obstructions to a lattice formulation of models with left-right asymmetry. In Chapter 5 we examine this problem and introduce a class of theories with left-right asymmetry which circumvent these difficulties by the use of the Higgs mechanism. We conclude with a candidate lattice action for the $SU(2) \times U(1)$ electroweak gauge theory which has all the features of the continuum model, including a low energy theory analogous to the Fermi theory.

PART I

INSTANTONS IN NON-ABELIAN GAUGE THEORIES

PART I

CHAPTER 1

INTRODUCTION TO INSTANTONS

1.1 Finite Action Solutions in Non-Abelian Gauge Theories

A finite action solution of the Euclidean SU(2) gauge theory was found in 1975 by Belavin et al. and subsequent work ('t Hooft, 1976; Callan et al., 1976) revealed a rich structure of the gauge theory vacuum. We begin with a brief review of the SU(2) gauge theory and examine the finite action solutions.

The action for the gauge theory is written in terms of vector potentials $A_\mu(x)$ where

$$A_\mu(x) = A^a_\mu(x) T^a \quad , \tag{1.1}$$

the $\{T^a\}$ being the generators of the gauge group. (The normalisation of the generators may be chosen at will. However for SU(2) we shall take $T^a = -\frac{i\sigma^a}{2}$, $\{\sigma^a\}$ being the Pauli matrices, wherefrom $\mathrm{Tr}\, T^a T^b = -\frac{1}{2}\delta^{ab}$). The field strength tensor is defined to be

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \ . \tag{1.2}$$

The pure gauge theory is then defined by the Euclidean action

$$S = \frac{-1}{2g^2} \int d^4x \ \mathrm{Tr}[F_{\mu\nu} F_{\mu\nu}] = \frac{1}{4g^2} \int d^4x \ F^a_{\mu\nu} F^a_{\mu\nu} . \tag{1.3}$$

General (position dependent) members of the gauge group $\Omega(x)$ may be constructed from the generators using arbitrary functions $\lambda^a(x)$

$$\Omega(x) \quad = \quad \exp[\lambda^a(x)T^a] \quad . \tag{1.4}$$

The action (1.3) has a local gauge invariance under the transformations

$$A_\mu \quad \rightarrow \quad \Omega\, A_\mu \Omega^{-1} + \Omega\, \partial_\mu \Omega^{-1} \tag{1.5}$$

from which

$$F_{\mu\nu} \quad \rightarrow \quad \Omega\, F_{\mu\nu}\, \Omega^{-1} \ . \tag{1.6}$$

If $F_{\mu\nu}$ vanishes then A_μ may be a general gauge transform from zero: that is

$$A_\mu \quad = \quad \Omega\, \partial_\mu\, \Omega^{-1} \tag{1.7}$$

for some Ω.

We now consider the finite action solutions. For a field configuration F to have finite action, we must have $F_{\mu\nu} \rightarrow 0$ as $r \rightarrow \infty$. (Strictly speaking, F must fall off faster than $^1/r^2$ as $r \rightarrow \infty$). Hence from (1.7)

$$A_\mu \quad \rightarrow \quad \Omega\, \partial_\mu\, \Omega^{-1} \qquad \text{as} \qquad r \rightarrow \infty$$

for some $\Omega(x)$. Now the domain of $\Omega(x)$ as $r \rightarrow \infty$ is the infinite three-sphere, S^3, so with every finite action field configuration there is associated a mapping of S^3 into the gauge group G. The homotopy characteristics of this map, $\pi_3(G)$ are well known and in particular

$$\pi_3(SU(2)) \quad = \quad Z \qquad . \tag{1.8}$$

That is, the mappings Ω fall into distinct homotopy classes labelled by the integers. This label is called the Pontrijagin index. (This index may be thought of as the number of times the three-sphere is wrapped around the parameter space of SU(2), [which is also S_3] and so is also called the winding number). The index is given by (Coleman, 1977)

$$\nu = \frac{1}{24\pi^2} \int d\theta_1\, d\theta_2\, d\theta_3\, \mathrm{Tr}[\epsilon_{ijk}\, \Omega(\partial_i\, \Omega^{-1}) \times$$

$$\Omega(\partial_j\, \Omega^{-1})\Omega(\partial_k\, \Omega^{-1})$$

where the θ_i are the angles which parameterize S_3. ν may also be expressed in terms of the fields (Belavin et al., 1975) as

$$\nu = \frac{1}{32\pi^2} \int d^4x\, F^a_{\mu\nu}\, \tilde{F}^a_{\mu\nu} \qquad (1.9)$$

where $\tilde{F}_{\mu\nu}$, the dual of F, is given by

$$\tilde{F}^a_{\mu\nu} = \tfrac{1}{2}\, \epsilon_{\mu\nu\sigma\rho}\, F^a_{\sigma\rho}$$

ν is also called the topological charge of the configuration. We have seen that all gauge configurations of finite action have an associated topological charge. It is not possible to deform continuously one such configuration to another of differing winding number while maintaining finiteness of the action.

All of the foregoing holds for any simple Lie group and in particular for SU(N) (Coleman, 1977; Bott, 1956). However, this is not the case for U(1) as $\pi_3(U(1))$ is trivial - hence there is no analogue of the topological charge in the Abelian gauge theory.

The first explicit finite action solution was discovered (for SU(2)) by Belavin et al. (1975) and given the name instanton. They used the inequality

$$\int (F - \tilde{F})^2\, d^4x \geq 0$$

which implies, (using $F^a_{\mu\nu} F^a_{\mu\nu} = \tilde{F}^a_{\mu\nu} \tilde{F}^a_{\mu\nu}$)

$$S[A] = \frac{1}{4g^2} \int F^a_{\mu\nu} F^a_{\mu\nu} d^4x \geqslant \frac{1}{4g^2} \int F^a_{\mu\nu} \tilde{F}^a_{\mu\nu} d^4x \quad .$$

Hence we have a lower bound on the action

$$S[A] \geqslant \frac{8\pi^2}{g^2} \nu \tag{1.11}$$

which attains its bound if and only if

$$F^a_{\mu\nu} = \pm \tilde{F}^a_{\mu\nu} \; . \tag{1.12}$$

Solutions of (1.12) are thus local minima of the action for a given winding number and so are solutions of the field equations. As mentioned in the prologue, they dominate the functional integral and so must be used as the centre points for any semi-classical expansions.

(1.12) is a first order equation and so is more tractable than the (second order) field equations. The solution for $\nu = 1$ (which is self dual) in the Landau gauge $(\partial_\mu A^\mu = 0)$ is

$$A^a_\mu(x) = 2 R^a{}_b \frac{\eta^b{}_{\mu\nu}(x-x_o)^\nu}{[(x-x_o)^2 + \rho^2]} \tag{1.13}$$

where $\eta^a{}_{\mu\nu}$ is given by

$$\eta^a{}_{\mu\nu} = \varepsilon_{oa\mu\nu} + \tfrac{1}{2} \varepsilon_{abc} \varepsilon_{bc\mu\nu} \quad . \tag{1.14}$$

(1.13) is localized (hence the name instanton) and contains four arbitrary parameters x^μ_o specifying the location of the instanton, thus reflecting the translation invariance of the theory. The arbitrary parameter ρ is the scale size of the instanton and reflects the scale invariance of the classical action. $R^a{}_b$ is a rotation matrix and describes the orientation of the instanton in colour space and hence, for SU(2), contains three arbitrary parameters. (This reflects the global gauge invariance retained in the theory).

From (1.13) we can derive $F_{\mu\nu}$ using (1.2) :-

$$F^a_{\mu\nu} = -4 R^a{}_b \frac{\eta^b{}_{\mu\nu}\, \rho^2}{[(x-x_o)^2 + \rho^2]^2} . \tag{1.15}$$

There is the corresponding anti-self-dual solution (for $\nu = -1$) called the anti-instanton, which is the same as (1.13) and (1.15) with η replaced by $\bar{\eta}$ where

$$\bar{\eta}^a{}_{\mu\nu} = \varepsilon_{oa\mu\nu} - \tfrac{1}{2}\, \varepsilon_{abc}\, \varepsilon_{bc\mu\nu} . \tag{1.16}$$

(Various properties of η and $\bar{\eta}$ are given in 't Hooft, 1976). Often we shall refer collectively to instantons and anti-instantons as instantons.

It may be established (Atiyah and Ward, 1977) that there are no other solutions of unit winding number. Higher ν solutions do exist (Jackiw et al., 1977) but we shall represent such solutions later by approximate solutions of superpositions of widely spaced instantons: the dilute gas approximation which we shall consider in a later section.

1.2 The Structure of the Vacuum

We have discovered that all field configurations of finite action have an associated topological charge: the classical vacua of the gauge theory also fall into discrete classes labelled by the same integer winding number n (Jackiw and Rebbi, 1976; Callan et al., 1976; Coleman, 1977). We shall denote these vacua by $|n>$. The naive perturbative vacuum $(A_\mu = 0)$ belongs to the sector with $n = 0$. Classically it would not be possible (in Minkowski space) to make continuous transitions from one sector to another but, as pointed out in the prologue, the instantons are solutions of the Euclidean field equations and so represent tunnelling solutions. Such tunnelling then makes possible transitions which would be forbidden classically. Field configurations with topological charge ν tunnel between vacua whose winding numbers differ by ν. The existence of this tunnelling means that the full quantum vacuum will in general be a superposition of $|n>$ vacua.

Consider the matrix element of the evolution operator between two vacua n and n'. We shall put the model in a large space-time box VT. The matrix element is

$$<n|e^{-HT}|n'> \quad = \quad N \int \mathcal{D}A_\mu \; e^{-S[A_\mu]} \tag{1.17}$$

where N is some normalization constant and the functional integral runs over fields which tunnel from $|n'>$ at $t = -T/2$ to $|n>$ at $t = +T/2$. We shall use the semi-classical approximation to evaluate (1.17) - the functional integral is dominated by the stationary points of S[A] which we shall describe by superpositions of well-separated instantons and anti-instantons. This is the dilute gas approximation: any interactions or overlapping of the instantons are neglected. This will be seen to be at least self-consistent.

First we need the expression for the fluctuation around a single instanton solution A_o. Put $A = A_o + \hat{A}$ so that

$$S[A] \quad = \quad S_o + \tfrac{1}{2}\int \hat{A}\, M\, \hat{A}\, d^4x + \ldots..$$

where $S_o = S[A_o]$, the first derivative of S at A_o vanishes as A_o is an extremum of the action, and

$$M\, \delta^4(x-y) \quad = \quad \left.\frac{\delta^2 S}{\delta A(x)\delta A(y)}\right|_{A=A_o} .$$

We make the g dependence explicit by recalling (1.3) and writing S/g^2 rather than S - (thus $S_o = 8\pi^2$ rather than $8\pi^2/g^2$). The fluctuations around a single instanton become

$$\int \mathcal{D}A\; e^{-S[A]/g^2} \quad = \quad e^{-S_o/g^2} \int \mathcal{D}\hat{A}\; e^{\frac{-1}{2g^2}\int \hat{A}M\hat{A}\, d^4x}$$

$$= \quad e^{-S_o/g^2}\, K \int d^4x \quad [1 + O(g^2)] \qquad (1.18)$$

where $K = (\det M)^{-\frac{1}{2}}$ and we have made the zero mode integral over space time explicit. (This step is examined more carefully in the next section.)

Consider now an arbitrary stationary point of $S[A]$ in (1.17), consisting of r instantons and s anti-instantons. Let us assign and order the time coordinates of the instantons

$$- T/2 < t_r < t_{r-1} < \ldots\ldots\ldots. < t_2 < t_1 < T/2.$$

The integral over space time for the instantons is

$$V^r \int_{-T/2}^{T/2} dt_1 \int_{-T/2}^{t_1} dt_2 \; \ldots\ldots \int_{-T/2}^{t_{r-1}} dt_r = \frac{V^r T^r}{r!} \; .$$

Similarly for the anti-instantons we have $\frac{V^s T^s}{s!}$ - we can distinguish between an instanton and an anti-instanton but not between two instantons or two anti-instantons. The determinant of gaussian fluctuations factorizes for large separations so we have

$$\langle n|e^{-HT}|n'\rangle$$

$$= \sum_r \frac{[VTK\, e^{-S_o/g^2}]^r}{r!} \sum_s \frac{[VTK\, e^{-S_o/g^2}]^s}{s!} \delta_{n-n',r-s} \; . \tag{1.19}$$

The δ-symbol ensures that the total topological charge of the configuration equals the change in winding number from $-T/2$ to $T/2$. Putting

$$\delta_{a,b} = \frac{1}{2\pi} \int_o^{2\pi} d\theta \; e^{i(a-b)\theta}$$

we obtain for the matrix element

$$\frac{1}{2\pi} \int_o^{2\pi} d\theta \; e^{i(n-n')\theta} \sum_r \frac{[VTK\, e^{-S_o/g^2 - i\theta}]^r}{r!} \sum_s \frac{[VTK\, e^{-S_o/g^2 + i\theta}]^s}{s!} \tag{1.20}$$

The dominant contribution comes from $r, s = KVT\, e^{-S_o/g^2 \mp i\theta}$ so the density of instantons, $\frac{r}{VT}$, is exponentially small. Hence they remain well separated, at any rate for small g, and the dilute gas approximation is at least self-consistent.

From (1.20) we have

$$\langle n|e^{-HT}|n'\rangle = \frac{1}{2\pi}\int_0^{2\pi} d\theta\ e^{i(n-n')\theta}\ \exp[2VTK\ e^{-S_o/g^2}\cos\theta] \qquad (1.21)$$

and so we find that the tunnelling amplitude is labelled by a continuous variable θ, $0 \leq \theta \leq 2\pi$.

Now the quantum tunnelling will mix the states $|n\rangle$ to generate eigenstates of the translation operator R, defined by

$$R|n\rangle = |n+1\rangle \ .$$

Because of translation invariance $[R, H] = 0$ and the orthonormality of the states $|n\rangle$ require R to be unitary. Hence its eigenvalues have unit modulus and we can call them $e^{-i\theta}$, corresponding to the state $|\theta\rangle$. Since the states $|n\rangle$ are complete, we may write

$$|\theta\rangle = \sum_n a_n|n\rangle$$

for some a_n. Application of R yields $a_n = e^{i\theta}\ a_{n-1}$. We may choose $a_o = 1$, since the $|\theta\rangle$ states are not normalisable, to give

$$|\theta\rangle = \sum_n e^{in\theta}|n\rangle, \qquad 0 \leq \theta \leq 2\pi \ . \qquad (1.22)$$

To find the energy of the θ vacuum we evaluate

$$\begin{aligned}\langle\theta|e^{-HT}|n\rangle &= \sum_m e^{-im\theta}\ \langle m|e^{-HT}|n\rangle \\ &= e^{-in\theta}\ \exp[2KVT\ \cos\theta\ e^{-S_o/g^2}]\end{aligned}$$

where we have used (1.21). However, since $|\theta\rangle$ is an eigenstate of H

$$\langle\theta|e^{-HT}|n\rangle = e^{-E(\theta)T} \cdot e^{-in\theta}$$

and so we can read off the energy density of the θ-vacuum (relative to the perturbative vacuum)

$$\frac{E(\theta)}{V} = -2K\cos\theta\, e^{-S_o/g^2} .$$

$\theta = 0$ appears to be the state of lowest energy density but it does not follow that this must be the true vacuum of the theory as the θ vacuum is stable under gauge-invariant perturbations. Imagine a gauge invariant operator P :- then $[P,H] = 0$ and the θ vacua must be eigenstates of P, eigenvalue p_θ say. Then if we evaluate the matrix element of P between two θ vacua we have

$$\begin{aligned}
\langle\theta'|P|\theta\rangle &= p_\theta\langle\theta'|\theta\rangle \\
&= p_\theta \sum_{n,n'} \langle n'|e^{-in'\theta'} e^{in\theta}|n\rangle \\
&= p_\theta \sum_n e^{in(\theta-\theta')} \\
&= 2\pi\, p_\theta\,\delta(\theta-\theta') .
\end{aligned}$$

We must regard θ as a new parameter in the theory. For $\theta \neq 0$, parity and time reversal symmetries are broken, so nature exhibits a state with $\theta \simeq 0$.

1.3 Instanton Density and the Dilute Gas Approximation

We conclude this chapter with a closer examination of the quantum fluctuations around a single instanton and remark on the implications this has on the dilute gas approximation. We can consider the general SU(N) model throughout this section. The calculation of the determinant of gaussian oscillations about a single instanton was first completed by 't Hooft (1976) and is rather lengthy: however the result can be derived up to a multiplicative constant quite easily (Coleman, 1977). We shall use dimensional regularization (the full calculation by this method is in Shore, 1979) and work in $d = 4 - \varepsilon$. The coupling constant g has dimension $[\text{mass}]^{\varepsilon/2}$. Now the only other dimensional parameter in the bare theory (in d dimensions) is the scale size of the instanton: hence the classical action becomes

$$S[A_o] = \frac{S_o(\varepsilon)\rho^{-\varepsilon}}{g^2}$$

where $S_o(0) = 8\pi^2$. In order to evaluate (1.18) we have to consider the zero modes carefully. There are d from space-time translations, 1 from scale transformations and from global SU(N) rotations there are $4N - 5$. This gives a total of $4N + d - 4$ zero modes. As each collective co-ordinate associated with each zero mode is pulled out there is the corresponding Jacobian which is $O(^1/g)$. (The Jacobian is $O(^1/g)$ because the classical solution A_o is also $O(^1/g)$ - see also 't Hooft, 1976, and Callan et al., 1978). Hence the dimensionless factor for each collective co-ordinate is $[\rho^{-\varepsilon}/g^2]^{\frac{1}{2}}$, giving a total factor

$$\int d\Omega \int d^d x_o \int \frac{d\rho}{\rho^{d+1}} \cdot \left[\frac{\rho^{-\varepsilon}}{g^2}\right]^{\frac{4N+d-4}{2}}$$

where $d\Omega$ is the element of integration in the gauge group and we have inserted $d+1$ powers of ρ in the denominator by dimensional analysis. Thus (1.18) becomes

$$C_b VT \int \frac{d\rho}{\rho^{d+1}} \left[\frac{\rho^{-\varepsilon}}{g^2}\right]^{\frac{4N+d-4}{2}} \cdot \exp\left[\frac{-S_o(\varepsilon)\rho^{-\varepsilon}}{g^2}\right] [1+O(g^2)] \qquad (1.23)$$

where C_b is a constant which is ultra-violet divergent and has the form (Shore, 1979)

$$C_b = C_R \exp\left[\frac{11N}{3}\cdot\frac{1}{\varepsilon}\right] \tag{1.24}$$

with C_R finite as $\varepsilon \to 0$.

We can remove the divergence by perturbative coupling constant renormalization (the short distance structure should not be affected by the instanton, which is an extended object). To one loop this is

$$\frac{\rho^{-\varepsilon}}{g^2} = \frac{1}{g_R^2(\rho^{-1})} + \frac{1}{\varepsilon}\,\frac{11N}{24\pi^2} \quad . \tag{1.25}$$

Collecting (1.24) and (1.25) together with (1.23) and letting $\varepsilon \to 0$ we have the one loop result

$$C_R VT \int \frac{d\rho}{\rho^5}\left[\frac{1}{g_R^2(\rho^{-1})}\right]^{2N} \exp\left[\frac{-8\pi^2}{g_R^2(\rho^{-1})}\right]\cdot[1 + O(g_R^2)] \ . \tag{1.26}$$

In the discussion of the last section, following (1.20), we established that in the dilute gas approximation the density of instantons per unit space-time volume is this one loop determinant factored by the space-time volume VT. Accordingly we may read off from (1.26) the density of instantons lying in the range ρ to $\rho+d\rho$

$$D(\rho)d\rho = C_R\left[\frac{1}{g_R^2(\rho^{-1})}\right]^{4N} \exp\left[\frac{-8\pi^2}{g_R^2(\rho^{-1})}\right]\frac{d\rho}{\rho^5} \tag{1.27}$$

The ρ-dependence in (1.27) may be made explicit by introducing an arbitrary momentum scale μ

$$\frac{1}{g^2} = \rho^{\varepsilon}\left[\frac{1}{g_R^2(\rho^{-1})} + \frac{1}{\varepsilon}\frac{11N}{24\pi^2}\right]$$

$$= \mu^{-\varepsilon}\left[\frac{1}{g_R^2(\mu)} + \frac{1}{\varepsilon}\frac{11N}{24\pi^2}\right]$$

so that

$$\frac{1}{g_R^2(\rho^{-1})} = \frac{1}{(\rho\mu)^{\varepsilon}}\left[\frac{1}{g_R^2(\mu)} + \frac{1}{\varepsilon}\frac{11N}{24\pi^2}\,[1-(\rho\mu)^{\varepsilon}]\right]$$

$$\rightarrow \frac{1}{g_R^2(\mu)} - \frac{11N}{24\pi^2}\,\ln(\rho\mu) \quad \text{as} \quad \varepsilon \to 0 \quad .$$

Applying this to (1.27) to one loop we have

$$D(\rho)d\rho = C_R\left[\frac{1}{g_R^2(\mu)}\right]\exp\left[\frac{-8\pi^2}{g_R^2(\mu)}\right](\rho\mu)^{11N/3}\frac{d\rho}{\rho^5} \quad . \tag{1.28}$$

We see from (1.28) that the density of instantons increases with increasing scale size (for all $N > 1$) : there will come a scale size when the instantons start overlapping and so the dilute gas approximation breaks down. Thus the infra-red divergence of the scale-size integral in (1.26) reflects the breakdown of the approximations and the onset of strong coupling physics for which the small g expansions (both perturbative and semi-classical) are invalid.

An approach frequently used (e.g. Callan et al., 1978) is to impose an upper cutoff Λ on the allowed scale sizes of the instantons by requiring the fraction of space-time filled by them to be some small arbitrary number such as, for example, 5%. Then Λ is determined by the equation

$$\int_0^\Lambda D(\rho) \frac{S_4 \rho^4}{2} d\rho = 0.05$$

(S_4 is the surface area of the unit three-sphere). This is clearly not a wholly satisfactory arrangement - fortunately there are some calculations which do not require an arbitrary cutoff. In this thesis we shall examine two situations with this general feature.

In Chapter 2 instantons are placed within instantons to examine their antiscreening effect and the range of integration is restricted by the outer scale size. The strong coupling problem is revealed when this outer scale size becomes too large and the approximations quickly break down.

In Chapter 3 the probability distribution function for the topological charge captured by a finite sphere is calculated and here only instantons having scale sizes in a precisely determined range can contribute to the calculation: consequently there is no need for an arbitrary cutoff.

CHAPTER 2

INSTANTONS WITHIN INSTANTONS AND COUPLING CONSTANT RENORMALISATION

2.1 Introduction and Motivation

We begin this chapter by reviewing a calculation of the interaction action of an instanton in a weak external field (Callan et al., 1978) and then give a more formal argument in which it is not necessary to make the weak field assumption. Typically, we shall be thinking of the external field being due to the field inside a larger instanton, the only condition being that the scale size of the outer instanton is large compared with that of the inner. The "external field" will then be constant over the scale size of the smaller instanton. The virtue of this treatment is that we shall be able to decorate small instantons (which do not produce a weak field within their scale size) with even smaller ones.

Callan et al. (1978) used their interaction action to derive the renormalization of the coupling constant associated with the external field due to the presence of a dilute gas of instantons within that field. This produces an effective coupling which deviates slowly from the perturbative one until a certain point is reached at which the deviation becomes much sharper. The coupling grows quickly with increasing scale size and the instantons rapidly become very dense. The dilute gas approximation then begins to break down and the expressions can be trusted no longer.

We shall extend this idea in a consistent way by requiring that the instantons themselves should be internally dressed in the same way. That is, at any scale size, dressed fields will be used to dress the fields at larger scale sizes.

An indirect motivation for this is due to the droplet model ideas of Kadanoff (1976) and Bruce and Wallace (1981). It was realized that a correct picture of the scaling behaviour of the droplets may only be achieved if, at all scale sizes, each phase is internally decorated with droplets of the opposite phase. The model thus constructed is clearly scale independent, as desired.

This 'self-consistent' instanton dressing gives a new effective coupling which reduces to the old one at small scale sizes since the effect due to small instantons is indeed slight. However, at larger scales, a threshold is reached, slightly before the one mentioned above. At this point the new effective coupling suddenly deviates very sharply from the perturbative coupling and the dilute gas approximation breaks down at once.

In this chapter we shall consider just the pure SU(2) gauge theory. Extension to the SU(3) gauge theory is direct and there is no qualitative difference (however we shall see in the next chapter that this is not always the case).

We shall need to superpose instantons to give (approximate) multi-instanton solutions - however the (SU(2)) Landau gauge instanton solution given in Chapter 1 (equations (1.13) and (1.15)) is not suited to this purpose. In this gauge the gauge field A_μ falls off as $^1/x$ outside the scale size of the instanton and the field strength $F_{\mu\nu}$ falls off as $^1/x^4$. If we add together two (well separated) solutions $\underline{A}^1_\mu$, A^2_μ

$$\underline{A}_\mu = \underline{A}^1_\mu + \underline{A}^2_\mu$$

and use the resulting gauge potential $\underset{\sim}{A}_\mu$ to construct $F_{\mu\nu}$ according to (1.2) we obtain

$$\underset{\sim}{F}_{\mu\nu} = \underset{\sim}{F}^1_{\mu\nu} + \underset{\sim}{F}^2_{\mu\nu} + \underset{\sim}{A}^1_\mu \times \underset{\sim}{A}^2_\nu + \underset{\sim}{A}^1_\nu \times \underset{\sim}{A}^2_\mu \quad . \tag{2.1}$$

Now the fields $\underset{\sim}{F}^1_{\mu\nu}$ and $\underset{\sim}{F}^2_{\mu\nu}$ both fall off as $^1/x^4$ as already remarked but the non-linear cross term falls off as $^1/x^2$ and hence dominates at large distances. However, if we use the singular gauge obtained from the Landau gauge by inversion (Bogomol'nyi and Fateev, 1977) we can escape this problem. In this gauge the instanton is

$$A^a_\mu = 2 R^a{}_b \, \bar{\eta}^b{}_{\mu\nu} \frac{(x-x_o)_\nu \, \rho^2}{[(x-x_o)^2 + \rho^2](x-x_o)^2} \tag{2.2}$$

from which

$$F^a_{\mu\nu} = 4 R^a{}_b \, \bar{\eta}^b{}_{\mu'\nu'} \, M_{\mu\mu'} \, M_{\nu\nu'} \frac{\rho^2}{[(x-x_o)^2 + \rho^2]^2} \tag{2.3}$$

where $M_{\mu\mu'} = \delta_{\mu\mu'} - 2 \hat{x}_\mu \hat{x}_{\mu'}$.

F still falls off as $^1/x^4$ but now A falls off as $^1/x^3$ - hence the cross-term in (2.1) falls off as $^1/x^6$.

In this gauge, $F_{\mu\nu}$ contains the anti-self-dual object $\bar{\eta}$ which we normally associate with anti-instantons. F must still be self-dual, however, and we can establish this as follows.

$$\begin{aligned} \tilde{F}^a_{\sigma\rho} &= \tfrac{1}{2} \, \varepsilon_{\sigma\rho\mu\nu} F^a_{\mu\nu} \\ &= 4 R^a{}_b \frac{\rho^2}{[(x-x_o)^2 + \rho^2]^2} \cdot \tfrac{1}{2} \, \varepsilon_{\sigma\rho\mu\nu} \, M_{\mu\mu'} \, M_{\nu\nu'} \, \bar{\eta}^b{}_{\mu'\nu'} \quad . \end{aligned} \tag{2.4}$$

Now

$$\varepsilon_{\sigma\rho\mu\nu}\, M_{\mu\mu'}\, M_{\nu\nu'}\, M_{\sigma\sigma'}\, M_{\rho\rho'} \;=\; (\det M)\cdot\, \varepsilon_{\sigma'\rho'\mu'\nu'}$$

and since $M_{\mu\nu}\, M_{\nu\lambda} \;=\; \delta_{\mu\lambda}$ we have

$$\varepsilon_{\sigma\rho\mu\nu}\, M_{\mu\mu'}\, M_{\nu\nu'} \;=\; (\det M)\cdot\, \varepsilon_{\sigma'\rho'\mu'\nu'}\, M_{\sigma'\sigma}\, M_{\rho'\rho}\,.$$

Hence

$$\begin{aligned}
&\tfrac{1}{2}\, \varepsilon_{\sigma\rho\mu\nu}\, M_{\mu\mu'}\, M_{\nu\nu'}\, \bar{\eta}^{b}_{\mu'\nu'} \\
&= (\det M)\cdot \tfrac{1}{2}\, \varepsilon_{\sigma'\rho'\mu'\nu'}\, \bar{\eta}^{b}_{\mu'\nu'}\cdot M_{\sigma\sigma'}\, M_{\rho\rho'} \\
&= -\,(\det M)\, M_{\sigma\sigma'}\, M_{\rho\rho'}\, \bar{\eta}^{b}{}_{\sigma'\rho'} \qquad (2.5)
\end{aligned}$$

Since $\bar{\eta}$ is anti-self-dual (see for example, 't Hooft, 1976). We can see easily that $\det M = -1$ by considering the eigenvalues of M:- construct four (independent) eigenvectors
(i) x_μ (eigenvalue -1); (ii) three mutually orthogonal vectors, all orthogonal to x_μ, $e^1{}_\mu$, $e^2{}_\mu$, $e^3{}_\mu$ (each has eigenvalue $+1$).
Then $\det M$ = (product of eigenvalues) = -1.

Using (2.5) and (2.4)

$$\begin{aligned}
\tilde{F}^{a}_{\sigma\rho} &= 4\, R^{a}{}_{b}\, \frac{\rho^2}{[(x-x_o)^2 + \rho^2]^2}\, M_{\sigma\sigma'}\, M_{\rho\rho'}\, \bar{\eta}^{b}_{\sigma'\rho'} \\
&= F^{a}_{\sigma\rho}
\end{aligned}$$

as required.

2.2 Interaction of an Instanton with a Weak External Field

We wish to examine this instanton solution (2.2), (2.3) in the presence of an external field, which for the moment will be taken to be weak. We will then want to calculate the interaction action S_{int}. Since instantons can appear inside an external field in any orientation within the group space, we will need to do the average over group directions by averaging over the matrices $R^a{}_b$ using the grand orthogonality theorem

$$\langle R^a{}_b \, R^c{}_d \rangle \quad = \quad \delta^{ac} \, \delta_{bd} \tag{2.6}$$

We shall then calculate $\langle \exp(-S_{int}) \rangle$.

The calculation of S_{int} proceeds as follows. We start with an instanton of scale size ρ at the origin, with gauge potential $\underline{A}^o_\mu$ and field strength $\underline{F}^o_{\mu\nu}$. The external field $\underline{F}^{ext}_{\mu\nu}$ is assumed to be weak and slowly varying over the length scale ρ. The gauge field for this is taken to be

$$\delta \underline{A}_\mu \quad = \quad - \tfrac{1}{2} \, \underline{F}^{ext}_{\mu\nu} \, x_\nu \tag{2.7}$$

and this is then treated as a perturbation around the instanton solution $\underline{A}^o_\mu$.

Now we take a sphere of radius R centre the origin, with $\rho << R$. R is chosen so that at $|x| = R$ the instanton field is small w.r.t. the external field. We calculate the action as follows

(i) Outside $|x| = R$ we may take $\underline{F}_{\mu\nu} = \underline{F}^{ext}_{\mu\nu}$.

(ii) Inside $|x| = R$ we take $\underline{F}_{\mu\nu}$ calculated from $\underline{A}^o_\mu + \delta\underline{A}_\mu$.

Near $|x| = R$, δA dominates and the small instanton is not important, while near $|x| = \rho$, A^o dominates and the external field is not important.

Therefore we may write

$$S = \frac{1}{g^2} \int_{|x|<R} \tfrac{1}{4} F^2 \, d^4x + \frac{1}{g^2} \int_{|x|>R} \tfrac{1}{4} F^2_{ext} \, d^4x$$

where in the first integral

$$\underline{F}_{\mu} = \partial_\mu [\underline{A}^o_\nu + \delta\underline{A}_\nu] - \partial_\nu [\underline{A}^o_\mu + \delta\underline{A}_\mu] + [\underline{A}^o_\mu + \delta\underline{A}_\mu] \times [\underline{A}^o_\nu + \delta\underline{A}_\nu]$$

$$= \underline{F}^o_{\mu\nu} + \underline{D}^o_\mu \, \delta\underline{A}_\nu - \underline{D}^o_\nu \, \delta\underline{A}_\mu + O(\delta\underline{A}^2) \qquad (2.8)$$

where, of course, $\underline{D}^o = (\partial_\mu + \underline{A}^o_\mu \times)$. We then drop the $O(\delta\underline{A}^2)$ term to get

$$\tfrac{1}{4}F^2 = \tfrac{1}{4}F^{o2} + \underline{F}^o_{\mu\nu} \cdot (\underline{D}^o_\mu \, \delta\underline{A}_\nu) + O(\delta\underline{A}^2) \quad . \qquad (2.9)$$

The 2nd order terms are dropped again. Thus we get

$$S_{int} = S - S_o = \frac{1}{g^2} \int_{|x|<R} \underline{F}^o_{\mu\nu} \cdot (\underline{D}^o_\mu \, \delta\underline{A}_\nu) d^4x \qquad (2.10)$$

$$= \frac{1}{g^2} \int_{|x|<R} \delta\underline{A}_\nu \cdot (\underline{D}^o_\mu \, \underline{F}^o_{\mu\nu}) d^4x + \frac{1}{g^2} \int_{|x|=R} d\Omega \, \hat{x}_\mu \, \underline{F}^o_{\mu\nu} \cdot \delta\underline{A}_\nu$$

upon integrating by parts. The first integral is zero because of the field equations and the second may be evaluated explicitly using the formulae for $\underline{F}^o_{\mu\nu}$ and $\delta\underline{A}_\mu$, (2.3) and (2.7). The answer is (see also Callan et al., 1978)

$$S_{int} = \frac{\rho^2\pi^2}{g^2} F^{a\ ext}_{\mu\nu} R^b{}_a \bar{\eta}_{b\mu\nu} . \qquad (2.11)$$

We now evaluate $\langle \exp(-S_{int}) \rangle$ using (2.6). To do this we have to expand the exponential to second order since $\langle R\rangle = 0$. We get

$$1 + \frac{\rho^4\pi^4}{2g^4} F^{a\ \text{ext}}_{\mu\nu} F^{c\ \text{ext}}_{\lambda\rho} \bar{\eta}_{b\mu\nu} \bar{\eta}_{d\lambda\rho} \langle R_a{}^b R_c{}^d\rangle + O(F^4_{ext})$$

$$= 1 + \frac{\rho^4\pi^4}{6g^4} \underline{F}^{ext}_{\mu\nu} \cdot \underline{F}^{ext}_{\lambda\rho} (\delta_{\mu\lambda}\delta_{\nu\rho} - \delta_{\mu\rho}\delta_{\nu\lambda} - \varepsilon_{\mu\nu\lambda\rho}) \tag{2.12}$$

where we have used the relation ('t Hooft, 1976)

$$\bar{\eta}_{b\mu\nu}\bar{\eta}_{b\lambda\rho} = \delta_{\mu\lambda}\delta_{\nu\rho} - \delta_{\mu\rho}\delta_{\nu\lambda} - \varepsilon_{\mu\nu\lambda\rho}$$

(2.12) then becomes

$$1 + \frac{\rho^4\pi^4}{6g^4} (F^{ext} - \tilde{F}^{ext})^2 + O(F^{ext})^4 \tag{2.13}$$

$$\simeq \exp\left[\frac{2\rho^4\pi^4}{3g^4} F^2_{ext}\right] + O(F^4_{ext}) \tag{2.14}$$

if F^{ext} is purely anti-self-dual (e.g. an anti-instanton).

The problem in this approach may now be seen. The final answer is $O(F^2_{ext})$: but several terms of this order have been neglected en route - e.g. the $O(\underline{\delta A}^2)$ terms in equations (2.8) and (2.9) - some of these could survive to <u>first</u> order in the expansion of $\langle\exp(-S_{int})\rangle$. There is also the question about how the calculation works when F_{ext} is not necessarily small - how does one cope with $O(F^4_{ext})$ for example?

Fortunately this approach may be saved when F_{ext} is not small. The important point to notice in the expansion in equation (2.14) is that it is really an expansion in $\rho^4F^2_{ext}$. Now if F_{ext} is the field inside an (anti)-instanton of scale size ρ' then $|F_{ext}| \sim \frac{1}{\rho'^2}$

and so we have an expansion in $\left[\rho/\rho'\right]^4$. Hence, if we require that F_{ext} varies only slowly over the scale size of the smaller instanton, then $\rho << \rho'$ and so we have a genuine expansion in terms of a small quantity. This is the approach followed in the next section. We also examine the second order terms and keep track of them carefully.

There is also the problem of whether we are dealing with instantons or anti-instantons at each stage. This is not a difficult point and is resolved later. For the moment, we shall take the small instanton to be self-dual and the external field to be anti-self-dual (anti-instanton).

2.3 The Interaction Action:- F_{ext} is not necessarily weak.

We use the same approach. Inside $|x| = \rho$ we have the instanton solution $\underline{F}^o_{\mu\nu}$, in $\rho < |x| < R$ we have the field given by $\underline{A}^o_\mu + \underline{\delta A}_\mu$ and outside $|x| = R$ we have $F = F_{ext}$. We have $\rho << R << \rho'$.

First, we must ask if the expression (2.7) is justified if F_{ext} is not necessarily small. Constructing the field tensor from the gauge field we have

$$\begin{aligned}
\underline{F}_{\mu\nu} &= \partial_\mu(\underline{\delta A}_\nu) - \partial_\nu(\underline{\delta A}_\mu) + \underline{\delta A}_\mu \times \underline{\delta A}_\nu \\
&= -\tfrac{1}{2}\,\underline{F}^{ext}_{\nu\mu} + \tfrac{1}{2}\,\underline{F}^{ext}_{\mu\nu} + \tfrac{1}{4}\,\underline{F}^{ext}_{\mu\lambda} \times \underline{F}^{ext}_{\nu\tau}\, x_\lambda x_\tau \\
&= \underline{F}^{ext}_{\mu\nu} + \tfrac{1}{4}\,\underline{F}^{ext}_{\mu\lambda} \times \underline{F}^{ext}_{\nu\tau}\, x_\lambda x_\tau \\
&\sim \frac{1}{\rho'^2} + \frac{1}{\rho'^4}\, x^2 \\
&= \frac{1}{\rho'^2}\left[1 + \frac{x^2}{\rho'^2}\right]
\end{aligned}$$

However, we are only using this $\underline{\delta A}_\mu$ in the range $\rho < |x| < R$, so the maximum attained by the non-Abelian part is at $|x| = R$ and this is $\frac{R^2}{\rho'^2} << 1$.

Hence the non-Abelian part may be dropped and we conclude that $\underline{\delta A}_\mu$ given by (2.7) may be used to describe the external field.

We now follow the construction of the previous section and calculate $\underline{F}_{\mu\nu}$ from $\underline{A}^o_\mu + \underline{\delta A}_\mu$, taking care with the second order terms. We get

$$\underline{F}_{\mu\nu} = \underline{F}^o_{\mu\nu} + \underline{D}^o_\mu \, \underline{\delta A}_\nu - \underline{D}^o_\nu \, \underline{\delta A}_\mu + \underline{\delta A}_\mu \times \underline{\delta A}_\nu \quad . \tag{2.15}$$

In fact the last term may be dropped with respect to the others when we consider squaring equation (2.15) as follows.

The largest cross-term involving the last term is

$$\underline{F}^o_{\mu\nu} \cdot \underline{\delta A}_\mu \times \underline{\delta A}_\nu \ .$$

Now when we calculate $\langle \exp(-S_{int}) \rangle$ by expanding the exponential and doing the group averages, this term will give zero to lowest order as it will contain only one R matrix.

The next largest term in (2.15) squared involving $\underline{\delta A}_\mu \times \underline{\delta A}_\nu$ is

$$[\underline{D}^o_\mu \, \underline{\delta A}_\nu] \cdot [\underline{\delta A}_\mu \times \underline{\delta A}_\nu] \quad \sim \quad 1/R \, [R/\rho'^2]^3$$

since $|\underline{\delta A}_\mu| \lesssim R/\rho'^2$.

This has to be integrated over all space which introduces a factor R^4: so this term behaves like

$$\sim \left[\frac{R}{\rho'}\right]^6 \quad . \tag{2.16}$$

Now the calculation due to Callan et al. gives, from (2.13)

$$<e^{-S_{int}}> \sim 1 + O[\rho^4/\rho'^4] \quad . \tag{2.17}$$

Hence (2.16) is an order of magnitude down on the old result and so may be dropped.

The last term in F^2, as given by (2.15), is

$$O(\underline{\delta A}^4) \quad \sim \quad \frac{R^4}{\rho'^8}$$

and so goes like $\left[\frac{R}{\rho'}\right]^8$ after integration which is smaller still. Similarly, all higher terms in expanding the exponential are at least as small as $(\delta A)^4$.

We now have F^2, having taken care with the second order terms, which may be dropped after all. We have

$$\begin{aligned}
\tfrac{1}{4}\underline{F}^2_{\mu\nu} &= \tfrac{1}{4}(\underline{F}^o_{\mu\nu})^2 + [\underline{D}^o_\mu \ \underline{\delta A}_\nu]\cdot\underline{F}^o_{\mu\nu} + \tfrac{1}{4}(\underline{D}^o_\mu \ \underline{\delta A}_\nu - \underline{D}^o_\nu \ \underline{\delta A}_\mu)^2 \\
&= \tfrac{1}{4}\underline{F}^{o2} + [\underline{D}^o_\mu \ \underline{\delta A}_\nu]\cdot\underline{F}^o_{\mu\nu} + \tfrac{1}{4}\underline{F}^2_{ext} + \underline{F}^{ext}_{\mu\nu}\cdot\underline{A}^o_\mu \times \underline{\delta A}_\nu \\
&\quad + \tfrac{1}{4}(\underline{A}^o_\mu \times \underline{\delta A}_\nu - \underline{A}^o_\nu \times \underline{\delta A}_\mu)^2
\end{aligned}$$

expanding the last term. We can now write

$$\begin{aligned}
S_{int} = \frac{1}{g^2}\int d^4x \Big[&[\underline{D}^o_\mu \ (\underline{\delta A}_\nu)]\cdot\underline{F}^o_{\mu\nu} + \underline{F}^{ext}_{\mu\nu}\cdot\underline{A}^o_\mu \times \underline{\delta A}_\nu \\
&+ \tfrac{1}{4}(\underline{A}^o_\mu \times \underline{\delta A}_\nu - \underline{A}^o_\nu \times \underline{\delta A}_\mu)^2\Big]
\end{aligned} \tag{2.18}$$

The first term in (2.18) is Callan et al.'s term as given in equation (2.10). As before this yields

$$\frac{2\pi^4}{3g^4}\ \rho^4 F^2_{ext} \tag{2.19}$$

as in equation (2.14), which is of order $[\rho/\rho']^4$. As in the case of justifying the use of δA, we will examine the other terms to show that they are of a higher order.

The second term in (2.18) contains just one $\underline{A}^o_\mu$ and hence only one matrix R. Hence, upon doing the group averages this term vanishes to lowest order in the expansion of $\langle \exp(-S_{int})\rangle$. The next highest term is the cross-term between the first and second terms of (2.18) which arises in the second order term in $\exp(-S_{int})$.

This is

$$\frac{\rho^2\pi^2}{g^2} F^{a\,ext}_{\mu\nu} R^a{}_b \bar{\eta}_{b\mu\nu} \frac{1}{g^2}\int \underline{F}^{ext}_{\mu\nu}\cdot\underline{A}^o_\mu \times \underline{\delta A}_\nu \, d^4x, \quad \text{where we have used (2.11)}$$

$$\sim \frac{\rho^2}{\rho'^2}\cdot R^4 \frac{1}{\rho'^2}\frac{\rho^2}{R^3}\cdot\frac{R}{\rho'^2} = \frac{\rho^4}{\rho'^4}\cdot\frac{R^2}{\rho'^2} << \frac{\rho^4}{\rho'^4} .$$

[Here we have used $|A^o{}_\mu| \sim \rho^2/R^3$].

Hence this term may be dropped, as it is much smaller than (2.19) - and so may all higher terms involving the second term in (2.18) as they are all at least as small.

Finally, we come to the third term in (2.18). First notice that it contains two R matrices and so will contribute to lowest order in exp $(-S_{int})$ - in the linear term. However, the higher order terms may be neglected as before - the largest of these is the cross-term between the first and the third term in (2.18), (arising from the quadratic term in $\exp(-S_{int})$). It is

$$\frac{\rho^2\pi^2}{g^2} F^{a\,ext}_{\mu\nu} R^a{}_b \bar{\eta}_{b\mu\nu}\frac{1}{g^2}\int \tfrac{1}{4}(\underline{A}^o_\mu \times \underline{\delta A}_\nu - \underline{A}^o_\nu \times \underline{\delta A}_\mu)^2 \, d^4x$$

$$\sim \frac{\rho^2}{\rho'^2}\cdot R^4 \cdot \left[\frac{\rho^2}{R^3}\;\frac{R}{\rho'^2}\right]^2 = \left(\frac{\rho}{\rho'}\right)^6 << \left(\frac{\rho}{\rho'}\right)^4 .$$

All higher terms are of this order or smaller and so may safely be dropped.

Now we can examine the lowest order term which contributes in the linear term in the exponential. It is

$$-\frac{1}{g^2}\int_{|x|<R} \tfrac{1}{4}(\underline{A}^o_\mu \times \underline{\delta A}_\nu - \underline{A}^o_\nu \times \underline{\delta A}_\mu)^2 \, d^4x \quad .$$

Using (2.2) and (2.7), this is straightforward to calculate; as was (2.11). The answer is

$$-\frac{1}{3}\frac{\pi^2}{g^2}\rho^4 F^2_{ext} \times \ell n \frac{R}{\rho} \quad .$$

At first sight, the appearance of a logarithm seems disconcerting. We must think about the relative magnitudes of ρ, R and ρ'. Recall first that R is fixed, with $\rho << R << \rho'$. Furthermore, we are thinking, at present, of small instantons, so that F_{ext} is large. Hence ρ' is small and so $\ell n\, R/\rho$ does not reflect an ultra-violet divergence, as R is not allowed to go off to infinity. We have

$$\begin{aligned} \langle e^{-S_{int}} \rangle &= 1 + \frac{2}{3}\frac{\pi^4}{g^4}\rho^4 F^2_{ext} - \frac{1}{3}\frac{\pi^2}{g^2}\rho^4 F^2_{ext} \times \ell n \frac{R}{\rho} \\ &= 1 + \frac{2}{3}\frac{\pi^4}{g^4}\rho^4 F^2_{ext}\left[1 - \frac{g^2}{2\pi^2} \times \ell n \frac{R}{\rho}\right] . \qquad (2.20) \end{aligned}$$

Comparing (2.20) with (2.14) we find a correction to Callan et al.'s terms. The correction factor is

$$1 - \frac{g^2}{2\pi^2}\ \ell n \frac{R}{\rho} \quad .$$

We need to estimate how large this correction factor may be. The coupling g refers to the coupling associated with the internal instanton. Using the lattice definition of the coupling, as adopted by Callan (1980) we have

$$g^2(\rho^{-1}) \quad = \quad \frac{12\ \pi^2}{11\ \ln \frac{1}{\Lambda_L \rho}} \qquad \text{(to one loop)} \qquad (2.21)$$

and so the correction factor is

$$\simeq \ 1 - \frac{6}{11} \ \frac{\ln R/\rho}{\ln \frac{1}{\Lambda_L \rho}} \qquad .$$

Now for all the instantons we will be considering, $\rho'\Lambda_L$ will remain small (< 0.01). This is because the instanton effects discussed in the next section have become important and the dilute gas has begun to break down by the time this scale size is reached.

Since $\rho << R << \rho' << \frac{1}{\Lambda_L}$, $\frac{6}{11} \ \frac{\ln R/\rho}{\ln \frac{1}{\Lambda_L \rho}}$ will remain small and so may be dropped.

The formula given by Callan et al. (1978) is then found to be correct even if the external field is not necessarily weak.

2.4 Instanton Renormalization of the Coupling Constant

The interaction action calculated in the last section may be absorbed into the action for F_{ext} and thought of as a renormalization of the coupling associated with the external field

$$\frac{1}{4g^2}\, F^2_{ext} - \frac{2}{3}\, \frac{\pi^4 \rho^4}{g^4}\, F^2_{ext} \quad = \quad \frac{1}{4g^2}\, F^2_{ext} \cdot \left[1 - \frac{8\pi^4}{3g^2}\, \rho^4\right] \quad .$$

However this was calculated using only one instanton and we know that we should rather populate the external field with a dilute gas of instantons, of density $D(\rho)d\rho$ per unit space-time for instantons with scale sizes in the range ρ to $\rho + d\rho$. $D(\rho)$ is derived in Chapter 1 and we may write it (see equation (1.27))

$$D(\rho)d\rho \quad = \quad C_2 \left[\frac{8\pi^2}{g^2(\rho^{-1})}\right]^4 \exp\left[\frac{8\pi^2}{g^2(\rho^{-1})}\right] \cdot \frac{d\rho}{\rho^5}$$

where C_2 is the appropriate constant for the SU(2) gauge group. It is a renormalization scheme dependent number and for the lattice definition of the coupling (2.21) it takes the value $(9.54)^{22/3}$ (see Callan, 1980 and Callan et al., 1979).

To incorporate the instanton gas we must integrate over scale sizes up to the external scale ρ', giving an effective coupling

$$\frac{F^2_{ext}}{4g^2_{eff}(\rho'^{-1})} = \frac{F^2_{ext}}{4g^2_{AF}(\rho'^{-1})}\left[1 - \int_0^{\rho'} d\rho\ D(\rho)\left[\frac{8\pi^2}{g^2_{AF}(\rho^{-1})}\right] \cdot \frac{\pi^2\rho^4}{3}\right] \tag{2.22}$$

where $g^2_{eff}(\rho^{-1})$ is the new effective coupling and $g^2_{AF}(\rho^{-1})$ is the (asymptotically free) perturbative one given by (2.21).

We may now generalize to any external field rather than a purely anti-self-dual one. Any field F may be split into self-dual and anti-self-dual parts, F_S and F_A, say

$$F_S = \frac{F + \tilde{F}}{2}, \qquad F_A = \frac{F - \tilde{F}}{2}.$$

Then $F = F_S + F_A$ and so $F^2 = F_S{}^2 + F_A{}^2$. Since the density function $D(\rho)$ is the same for both instantons and anti-instantons, the renormalization of the field F is the same for its self-dual and anti-self-dual parts - hence (2.22) is true for any F_{ext}.

We may then write

$$\frac{1}{g^2_{eff}(\rho'^{-1})} = \frac{1}{g^2_{AF}(\rho'^{-1})}\left[1 - \int_0^{\rho'} d\rho\, D(\rho)\, \frac{8\pi^2}{g^2(\rho^{-1})} \cdot \frac{\pi^2\rho^4}{3}\right] \tag{2.23}$$

or in differential equation form

$$d\left[\frac{g^2{}_{AF}}{g^2_{eff}}\right] = -\frac{C_2\pi^2}{3}\left[\frac{8\pi^2}{g^2(\rho^{-1})}\right]^5 \exp\left[\frac{-8\pi^2}{g^2(\rho^{-1})}\right]\frac{d\rho}{\rho} \ . \tag{2.24}$$

The boundary condition for (2.24) is $g^2_{AF}/g^2_{eff} = 1$ at $\rho = 0$: this is clear from (2.23).

In the framework presented by Callan et al. (1978 and 1979) $g^2(\rho^{-1})$ on the right hand side of (2.23) and (2.24) is understood as being the perturbative coupling $g^2_{AF}(\rho^{-1})$. (2.24) then describes the deviation of g_{eff} from g_{AF} with increasing scale size. The solution to this equation is tabulated in Table 2.1. Again we are using the conventions as adopted by Callan (1980) in which $g^2_{AF}(\rho^{-1})$ is given by the right hand side of (2.21).

We see that g^2_{eff} is the same as g^2_{AF} until $\rho\Lambda_L \simeq 0.007$. At this point the right hand side of (2.24) becomes significantly non-zero and a slow departure takes place - g^2_{eff} increasing faster than g^2_{AF}. This departure becomes increasingly rapid until at $\rho\Lambda_L \simeq 0.011$ it has become extreme and the values are not to be trusted due to the breakdown of the diluteness of the gas.

We now examine the effect of internally decorating the instantons with smaller, dressed instantons. If we look at a given scale size

TABLE 2.1

g_{eff}^2 calculated from $d\left[\frac{g_{AF}^2}{g_{eff}^2}\right] = -\frac{\pi^2 C_2}{3}\left[\frac{8\pi^2}{g_{AF}^2}\right]^5 \exp\left[-\frac{8\pi^2}{g_{AF}^2}\right]\frac{d\rho}{\rho}$

using (2.21) for g_{AF}^2

$\rho\Lambda_L$	g_{AF}^2	g_{eff}^2
0	0	0
0.001	1.56	1.56
0.005	2.03	2.04
0.007	2.17	2.26
0.009	2.29	2.80
0.010	2.34	3.63
0.011	2.39	6.78
0.0113	2.40	10.3
0.0115	2.41	16.7
0.0117	2.42	51.5
0.01175	2.42	114

this dressing is carried out simply by using the coupling of the dressed fields g_{eff} to describe the instantons which are to be used to dress the fields at that scale size. The instantons used to decorate the field will then be already dressed themselves. This is done by identifying $g^2(\rho^{-1})$ on the right hand side of (2.24) with $g^2_{eff}(\rho^{-1})$. The boundary condition is preserved the same because $g^2_{eff} = g^2_{AF} = 0$ at $\rho = 0$.

The solution to this new equation is tabulated in Table 2.2. Since the effect due to small instantons is very slight, $g^2_{eff} = g^2_{AF}$ for small scale sizes. However, when the right hand side of (2.24) becomes important, g^2_{eff} rises very much faster than previously. The anti-screening effect produced by the smaller instantons enhances the change in g^2_{eff}, so the right hand side of (2.24) rises sharply and g^2_{eff} itself grows very quickly. As before, the instantons then become very dense and we have to stop calculating. Qualitatively this coupling has the same features as the old effective coupling, but the effect is sharpened and takes place slightly earlier. The threshold at which deviation from g^2_{AF} takes place is $\rho\Lambda_L \simeq 0.0068$.

TABLE 2.2

g^2_{eff} calculated from $d\left[\frac{g^2_{AF}}{g^2_{eff}}\right] = -\frac{\pi^2 C_2}{3}\left[\frac{8\pi^2}{g^2_{eff}}\right]^5 \exp\left[-\frac{8\pi^2}{g^2_{eff}}\right]\frac{d\rho}{\rho}$

using (2.21) for g^2_{AF}

$\rho\Lambda_L$	g^2_{AF}	g^2_{eff}
0	0	0
0.001	1.56	1.56
0.002	1.73	1.73
0.003	1.85	1.85
0.004	1.95	1.95
0.005	2.03	2.04
0.006	2.11	2.15
0.0065	2.14	2.24
0.0066	2.14	2.27
0.0067	2.15	2.33
0.0068	2.16	2.56
0.00685	2.16	5340

2.5 Conclusion

We have extended the treatment developed by Callan et al. (1978) for the effective action for an instanton in a weak background field to an approach in which the external field is not necessarily weak, so long as it is slowly varying over the scale size of the instanton and small compared to the field at the centre of the instanton. This result is necessary if we are to dress small instantons with even smaller ones.

We discussed the effective coupling produced by the presence of a gas of instantons within the external field. If the instantons used to decorate the field are themselves dressed, then we have a self-consistent picture of this dressing process. The effective coupling so produced is similar to the previous effective coupling but it rises much more steeply in the crossover to strong coupling.

It is thought that this instanton effect may be responsible for the crossover from weak to strong coupling that occurs in QCD. Plotted in Figure 2.1 are the relevant $\hat{\beta}$-functions:- $\hat{\beta}$ is defined by

$$\hat{\beta}(g) = \frac{d\ \ell ng}{d\ \ell n\rho} = \frac{1}{g}\ \beta(g) \quad .$$

Now the strong coupling is expected to behave as (see, for example, Callan 1980)

$$g^2(\rho^{-1}) \sim e^{-\sigma\rho^2}$$

and so $\hat{\beta}_{strong}(g) = \ell ng^2$.

(σ is the string tension, which cancels out of $\hat{\beta}$).

The asymptotic freedom $\hat{\beta}$-function is also known analytically - with our conventions it is $\hat{\beta}_{AF}(g) = \frac{11}{24\pi^2} g^2$.

These two $\hat{\beta}$-functions are plotted in Figure 2.1 together with the $\hat{\beta}$-functions of both the old and the new effective coupling, which may be obtained by rearranging (2.24)

$$\hat{\beta}_{eff}^{(old)}(g_{eff}^{o}) = \hat{\beta}_{AF}(g_{AF}) + \frac{C_2\pi^2}{6}\left[\frac{g_{eff}^{o}}{g_{AF}}\right]^2\left[\frac{8\pi^2}{g_{AF}^2}\right]^5 \exp\left(\frac{-8\pi^2}{g_{AF}^2}\right)$$

and

$$\hat{\beta}_{eff}^{(new)}(g_{eff}^{N}) = \hat{\beta}_{AF}(g_{AF}) + \frac{C_2\pi^2}{6}\left[\frac{g_{eff}^{N}}{g_{AF}}\right]^2\left[\frac{8\pi^2}{(g_{eff}^{N})^2}\right]^5 \exp\left(\frac{-8\pi^2}{(g_{eff}^{N})^2}\right)$$

where g_{eff}^{o} and g_{eff}^{N} refer to the old and new effective coupling respectively. We can see from Figure 2.1 that the new effective coupling serves only to take the theory into the strong coupling regime at a slightly earlier stage than occurs with the old effective coupling.

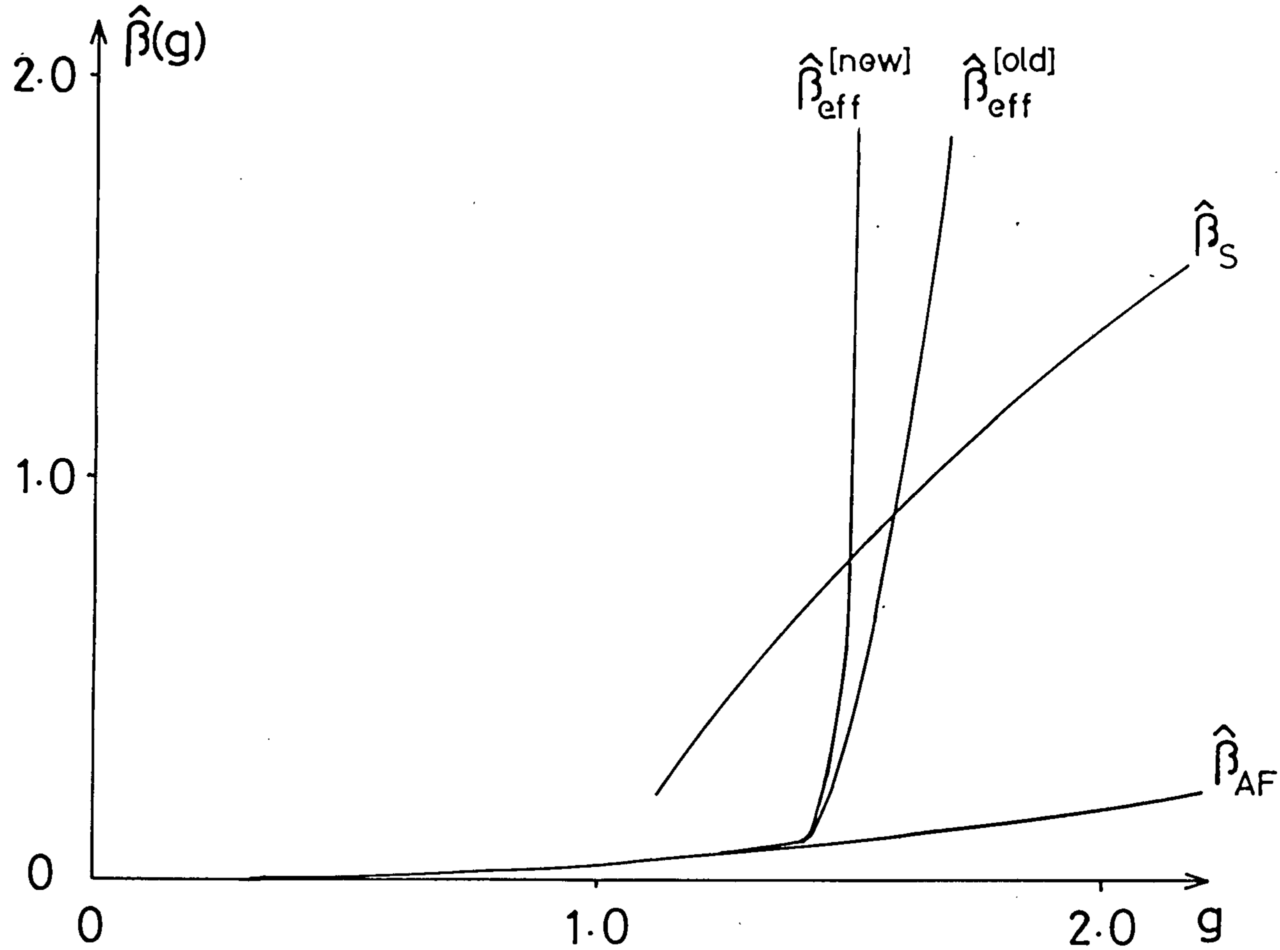


Figure 2.1

$\hat{\beta}$ versus g for the perturbative coupling ($\hat{\beta}_{AF}$) the strong coupling ($\hat{\beta}_s$) and the old and the new effective couplings induced by the instanton gas. The new effective coupling brings the theory into the strong coupling regime at a slightly earlier stage.

CHAPTER 3

DISTRIBUTION OF TOPOLOGICAL CHARGE

3.1 Topological Charge Inside a Sphere

We saw in Chapter 1 that all finite action field configurations have an associated topological charge as defined by equation (1.9). Since this is the integral of a local quantity we may define a topological charge density

$$\frac{1}{32\pi^2} F^a_{\mu\nu} \tilde{F}^a_{\mu\nu} \quad . \tag{3.1}$$

Note that (3.1) is gauge invariant.

In the dilute gas approximation we typically have a configuration of widely spaced instantons and anti-instantons. Imagine that we place a sphere of radius R in this configuration and measure the topological charge 'captured' by it, defined by

$$q = \frac{1}{32\pi^2} \int_{\text{sphere}} F^a_{\mu\nu} \tilde{F}^a_{\mu\nu} \, d^4x \tag{3.2}$$

We could repeat this process many times and so build up a probability distribution function for the topological charge contained in the sphere: it is this probability distribution function which we shall calculate in this chapter.

We begin by calculating the topological charge contained in a sphere of radius R and centred on the origin due to a single instanton of scale size ρ located at σ_μ (which need not be inside the sphere). From (1.15) we have the field strength tensor (in the Landau gauge)

$$F^a_{\mu\nu} = -4 R^a_{\ b} \eta^b_{\ \mu\nu} \frac{\rho^2}{[(x-\sigma)^2 + \rho^2]^2}$$

$R^a_{\ b}$ is orthogonal and so satisfies $R^a_{\ b} R^a_{\ c} = \delta_{bc}$. Hence (3.1) becomes

$$\frac{1}{2\pi^2} \eta^b_{\ \mu\nu} \eta^b_{\ \mu\nu} \frac{\rho^4}{[(x-\sigma)^2 + \rho^2]^4} .$$

Using the relation ('t Hooft, 1976)

$$\eta^b_{\ \mu\nu} \eta^b_{\ \mu\nu} = 12$$

we obtain

$$q = \frac{6\rho^4}{\pi^2} \int \frac{d^4x}{[(x-\sigma)^2 + \rho^2]^4} .$$

This integral may be evaluated analytically. Going to spherical polar coordinates and choosing the Z direction along σ_μ we have

$$d^4x = r^3 dr \sin^2\theta \, d\theta \sin\psi \, d\psi \, d\phi$$

with $0 \leqslant \theta, \psi \leqslant \pi$ and $0 \leqslant \phi \leqslant 2\pi$.

and so, doing the azimuthal integrals, (3.3) becomes

$$q = \frac{24\rho^4}{\pi} \int_{r=0}^{R} r^3 dr \int_{\theta=0}^{\pi} d\theta \frac{\sin^2\theta}{(r^2+\rho^2+\sigma^2-2r\sigma\cos\theta)^4}$$

$$= \frac{24}{\pi} \int_0^{R/\rho} ds \frac{s^3}{(2s\lambda)^4} \cdot I(s,\lambda) \tag{3.4}$$

where we have rescaled the variables as follows:-

$$s = r/\rho \quad , \qquad \lambda = \sigma/\rho$$

and

$$I(s,\lambda) = \int_0^{\pi} \frac{\sin^2\theta \, d\theta}{[c - \cos\theta]^4} \tag{3.5}$$

where $$c = \frac{s^2 + \lambda^2 + 1}{2s\lambda}$$

- note that $c > 1$ for all s and λ. (3.5) may be evaluated by using the standard substitution $\tan\frac{\theta}{2} = t$ which yields

$$I(s,\lambda) = \frac{1}{(c^2-1)^{5/2}} [(c-1)I_1 + 2 I_2]$$

where $$I_1 = 8\int_0^{\infty} \frac{\tau^2 \, d\tau}{[1+\tau^2]^3} = \frac{\pi}{2}$$

and $$I_2 = 8\int_0^{\infty} \frac{\tau^2 \, d\tau}{[1+\tau^2]^4} = \frac{\pi}{4} \quad .$$

Hence

$$I(s,\lambda) = \frac{\pi c}{2(c^2-1)^{5/2}} \quad .$$

Returning to (3.4) we now have

$$q = \frac{3}{4\lambda^4} \int_0^{R/\rho} \frac{ds}{s} \frac{c}{(c^2-1)^{5/2}}$$

$$= 12 \int_0^{R/\rho} \frac{(s^2 + \lambda^2 + 1)s^3 \, ds}{[s^4+2s^2(1-\lambda^2)+(\lambda^2+1)^2]^{5/2}} \quad .$$

This final radial integral may be done by completing the square in the denominator

$$s^4 + 2s^2(1-\lambda^2) + (\lambda^2+1)^2 \quad = \quad [s^2 + (1-\lambda^2)]^2 + 4\lambda^2$$

and so using the substitution $s^2 + (1-\lambda^2) = 2\lambda \tan\theta$. This yields several elementary trigonometric integrals. After considerable simplification we arrive at the final answer (having transformed back to the original variables)

$$q \quad = \quad \frac{1}{2} + \frac{1}{2}\,[4\rho^2\sigma^2 + (R^2 + \rho^2 - \sigma^2)^2]^{-3/2}$$

$$\times[(R^2+\rho^2-\sigma^2)^3 - 2\rho^6 - 6\rho^4R^2 + 6\rho^2R^2\sigma^2 - 6\rho^2\sigma^4] \ . \qquad (3.6)$$

We can easily check that at $\rho = 0$ we get $q = 1$ for $\sigma < R$ and $q = 0$ for $\sigma > R$, just as expected. The case $\sigma = R$ needs slightly more care. We need to put $\sigma = R$ first and then let $\rho \to 0$: this yields $q = \frac{1}{2}$, also as expected. Also, for $\sigma \to 0$ we obtain

$$q \quad = \quad 1 \ - \ \frac{\rho^6 + 3\rho^4\ R^2}{(R^2 + \rho^2)^3} \qquad (3.7)$$

which can easily be checked by evaluating (3.3) directly with $\sigma = 0$.

To illustrate (3.6) more easily we have plotted q against σ/R for various values of ρ in Figure 3.1. (Note that q is dimensionless so instead of thinking of q as $q(R, \rho, \sigma)$ we can think of it as $q(\rho/R, \sigma/R)$. In the next section we shall rescale ρ and σ in just this way). For the most part we shall consider just the range $\frac{1}{2} \leqslant q \leqslant 1$ and this is the range in which we shall calculate the probability density function. From Figure 3.1 and equation (3.6) we see that for $q \geqslant \frac{1}{2}$ we must have $\sigma \leqslant R$. (Alternatively one could recall that the topological charge distribution is spherically symmetric

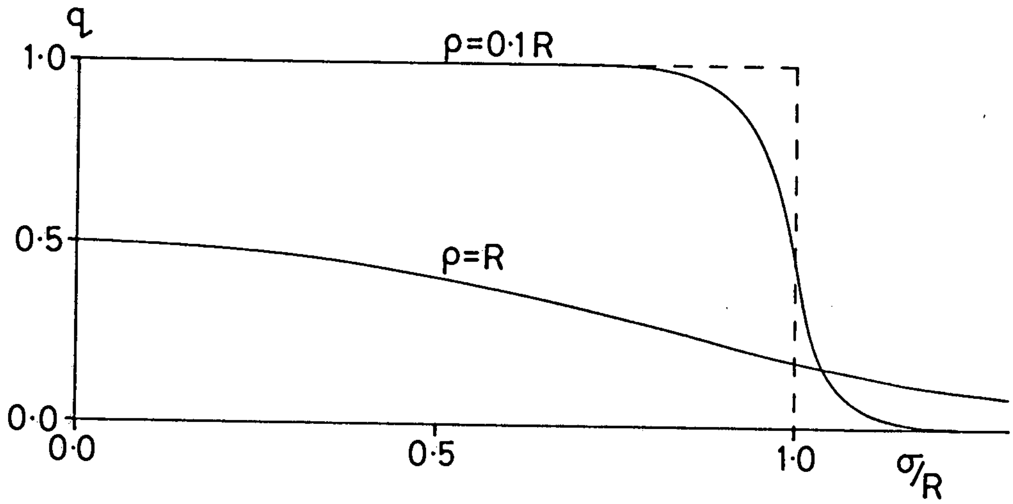


Figure 3.1

The variation of topological charge with instanton location for instantons of various scale sizes. The broken line shows the $\rho \to 0$ limit.

about the instanton centre: consequently it is impossible for the sphere to capture half or more of the topological charge if the instanton lies outside the sphere).

Furthermore, for $\frac{1}{2} \leq q \leq 1$ we must also have $\rho \leq R$. To see this, set $R = \rho$ in (3.7) - we obtain $q = \frac{1}{2}$ (showing that exactly half of the topological charge of an instanton is found within its own scale size). Now q decreases as σ is increased (for fixed ρ) and so the maximum value of q is at $\sigma = 0$. Hence to have $q > \frac{1}{2}$ we must have $\rho < R$. (We can also conclude this by inspection of Figure 3.1).

3.2 The Probability Distribution Function for Topological Charge

We are now in a position to calculate the probability density function $P(q)dq$ that a measurement of the total topological charge q contained in a sphere of radius R lies in the range q to $q + dq$, for $\frac{1}{2} \leq q \leq 1$. We will assume that the dominant contribution comes from single instanton effects. This is at least a consistent approach because the instanton scale size ρ must be less than the sphere radius R: consequently if we take R to be somewhat less than the hadron scale size then we know that the gas of such instantons will be dilute. Thus we work self-consistently within the dilute gas approximation. We note also that anti-instantons contribute to $P(q)$ for $q < 0$ and that $P(q)$ is symmetric about $q = 0$.

We already have an expression for the density $D(\rho)d\rho$ of instantons with scale sizes in the range ρ to $\rho + d\rho$ (1.28).

For given q, then, the contribution to the distribution function $P(q)dq$ due to instantons lying in the spherical shell $[\sigma, \sigma+d\sigma]$ is

$$D(\rho)d\rho \cdot S_4\sigma^3 \, d\sigma \tag{3.8}$$

where S_4 is the surface area of the unit three-sphere. In (3.8) ρ is determined uniquely by q and σ, so we shall have to invert (3.6) to find $\rho = \rho(q, \sigma, R)$. Integrating over all the values of σ which can contribute (recalling $\sigma \lesssim R$ from the previous section) we obtain

$$P(q) = S_4 \int_0^R d\sigma \, \sigma^3 \, D(\rho) \left|\frac{\partial\rho}{\partial q}\right| \tag{3.9}$$

with $\rho = \rho(q, \sigma, R)$. (Note that we cannot normalize $P(q)$ so that $\int_{-\infty}^{\infty} P(q)dq = 1$ as we do not calculate all of the distribution). Equation (3.9) may be derived directly from the functional integral for the gauge theory. The partition function is

$$Z = \int \mathcal{D}A_\mu \, e^{-S[A]} .$$

Inserting the identity

$$1 = \int_{-\infty}^{\infty} dq \, \delta(q - Q[A])$$

where $Q[A]$ is defined to be the right hand side of (3.2), we have

$$Z = \int_{-\infty}^{\infty} dq \, Z(q)$$

where

$$Z(q) = \int \mathcal{D}A_\mu \, e^{-S[A]} \, \delta(q - Q[A]) . \tag{3.10}$$

We can interpret $Z(q)$ as being the required distribution function $P(q)$ up to an overall constant. We may calculate (3.10) by a semi-classical expansion about the appropriate classical minimum, A_c, of the action. This will be a single instanton, as the zero instanton configuration does not satisfy the constraint. To lowest order in g^2, $Q[A] = Q[A_c](1 + O(g^2))$ and of course $Q[A_c] = Q(R, \rho, \sigma)$ as defined by equation (3.6). The gaussian fluctuations around the instanton have been considered in Chapter 1. The fluctuations orthogonal to the collective co-ordinates are not affected by the constraint: accordingly we find

$$Z(q) = \int d^4x \int D(\rho)d\rho \; \delta(q - Q(R, \rho, \sigma)) \; .$$

Now $$\int d^4x = S_4 \int \sigma^3 \, d\sigma$$

and $$\int D(\rho)d\rho \cdot \delta(q - Q(R, \rho, \sigma)) = D(\rho) \frac{\partial\rho}{\partial q}$$

with $\rho = \rho(q, \sigma, R)$, so we have

$$Z(q) = S_4 \int \sigma^3 \, D(\rho) \left|\frac{\partial\rho}{\partial q}\right| d\sigma$$

as before (see equation (3.9)).

Substituting (1.28) for $D(\rho)$ and collecting all the constants together as C we obtain

$$\begin{aligned} P(q) &= C \int_0^R \sigma^3 (\rho\mu)^{\frac{11N}{3}} \rho^{-5} \left|\frac{\partial\rho}{\partial q}\right| d\sigma \\ &= C(R\mu)^{\frac{11N}{3}} \, n(q) \end{aligned} \tag{3.11}$$

where

$$n(q) = \int_0^1 \hat{\sigma}^3 \, \hat{\rho}^{\frac{11N}{3} - 5} \left| \frac{\partial \hat{\rho}}{\partial q} \right| d\hat{\sigma} \qquad (3.12)$$

where $\hat{\rho} = \hat{\rho}(q, \hat{\sigma})$ and we have rescaled to dimensionless variables $\hat{\sigma} = \sigma/R$, $\hat{\rho} = \rho/R$. We keep the N of SU(N) explicit as we shall find a marked difference between the behaviour for SU(2) and SU(3).

We cannot invert (3.6) analytically and so must resort to numerical methods in order to calculate (3.12). A numerical integration routine was used together with a subroutine to solve (3.6) for $\hat{\rho}$ each time a value for the integrand of (3.12) was required. In Figure 3.2 we have plotted $n(q)$ for $\frac{1}{2} \leqslant q \leqslant 1$ for both SU(2) and SU(3) - and it is clear that the behaviour near $q = 1$ is completely different for SU(2) and SU(3).

In the region near $q = 1$ we can find the approximate form for P(q) analytically. Since q is close to 1, $\hat{\rho}$ must be very small and so we can expand (3.6) in powers of $\hat{\rho}$. The result is

$$q = 1 - \hat{\rho}^4 \cdot f(\hat{\sigma}) + O(\hat{\rho}^6)$$

which we can invert to give

$$\hat{\rho} = (1 - q)^{\frac{1}{4}} f^{-\frac{1}{4}} (\hat{\sigma}) \;\; [1 + O(1 - q)^{\frac{1}{2}}] \; .$$

Substituting this into (3.12) we find the approximate q-dependence of P(q) near $q = 1$:-

$$P(q) \sim (1 - q)^{\frac{11N}{12} - 2} \; . \qquad (3.13)$$

Thus for SU(2), P(q) goes to infinity as $(1 - q)^{-1/6}$ as $q \to 1$, while for SU(3), P(q) goes to zero as $(1 - q)^{3/4}$. These analytical

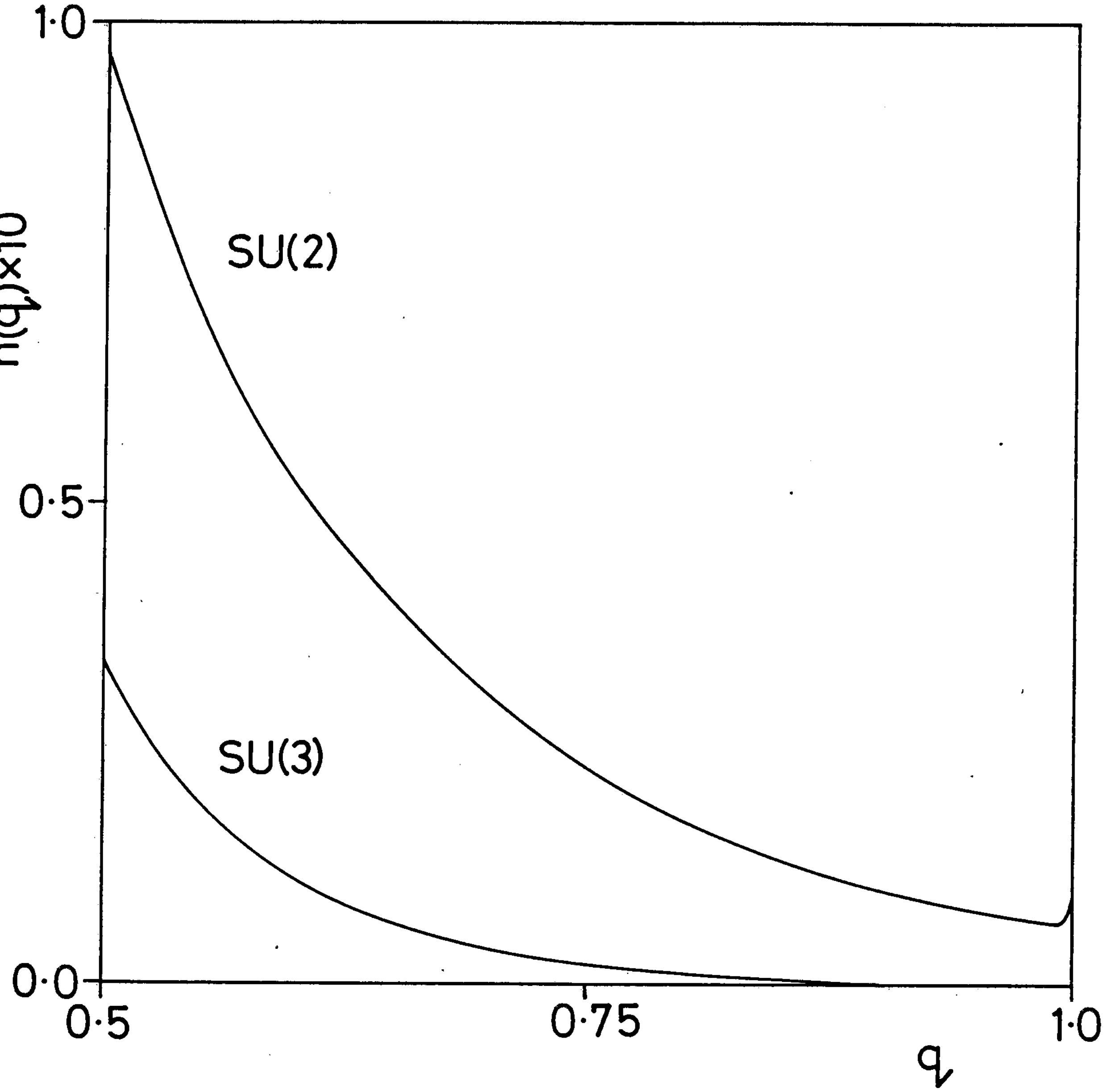


Figure 3.2

Distribution of topological charge in SU(2) and SU(3) gauge theories.

results fit well to the numerical data for $(1-q) \lesssim 10^{-4}$, which corresponds to $\hat{\rho} \lesssim \frac{1}{10}$.

3.3 Effects of a Space-time Lattice and Comparison with Monte Carlo Results

The results of the previous section may not be verified experimentally but they can be compared with computer Monte Carlo simulations of the gauge field defined on the lattice, as described in the prologue and in Chapter 4. Some studies searching for instantons in this way have already been performed (Ishikawa et al., 1983) in which the topological charge contained in a finite box was measured and corresponding frequency histograms plotted. Before examining these results, however, we shall describe how the lattice might affect the calculations of the previous section.

If the lattice gauge theory is to mimic successfully the continuum physics we would expect it to describe instantons of a length scale larger than the lattice spacing, a, but we would not expect to see instantons with a scale size much smaller than a. The simplest way to take this into account is to impose some minimum scale size ρ_{min} on the instantons. Now for a given value of q, ρ decreases as σ is increased (see Figure 3.1) and so a minimum value of ρ implies a maximum value of σ, given by

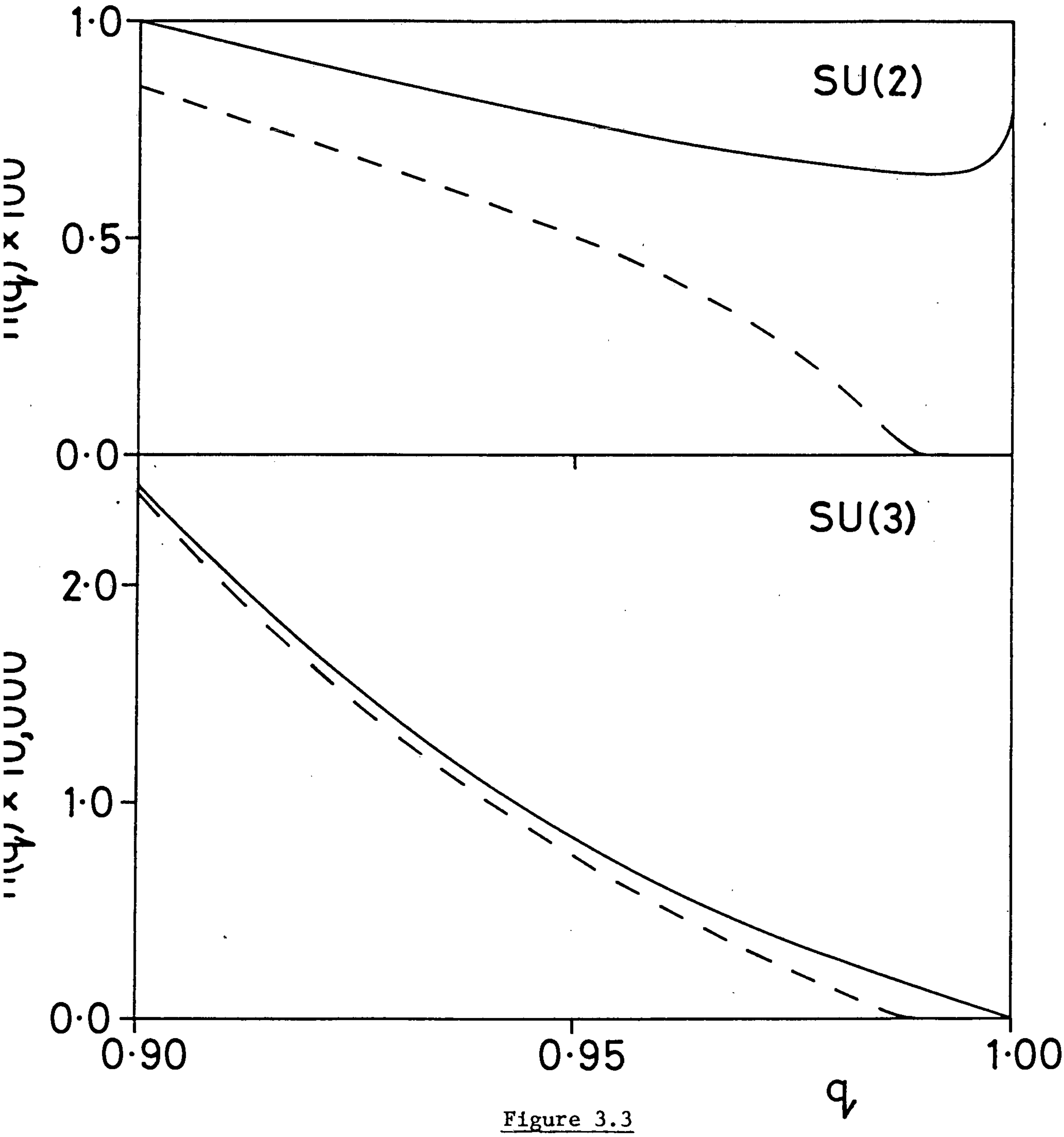
$$\sigma_{max} = \sigma(\rho_{min}, q, R)$$

where $\sigma(\rho, q, R)$ is found by inverting (3.6). We know that for $\frac{1}{2} \lesssim q \lesssim R$ we have $\sigma < R$ so the only change is that the upper limit of the integral in (3.12) is replaced by $\hat{\sigma}_{max}$. The integral must be re-evaluated and we have done this for $\rho_{min} = \frac{a}{2}$ and $R = 2a$ (which corresponds to representing the sphere by a 4^4 lattice, in line with the Monte Carlo studies we shall shortly discuss). The results are changed significantly only in the region of q near 1 since this is where small instantons - assumed to be absent on the lattice - dominate. These results are plotted in Figure 3.3 for $0.9 \leqslant q \leqslant 1$, together with the original continuum results. For a given ρ, there is a maximum value of q, occurring at $\sigma = 0$. Since q increases as ρ decreases (for given σ - in particular for $\sigma = 0$ - see eq. (3.7)) there is a maximum value of q corresponding to $\rho = \rho_{min}$ and $\sigma = 0$ given by $q_{max} = q(\rho_{min}, 0, R)$. Consequently the single instanton contribution to $P(q)$ must vanish for $q > q_{max}$ and the dominant contribution to $P(q)$ must come from two instanton effects. For $\rho_{min} = a/2 = R/4$ we have, from (3.7), $q_{max} \simeq 0.990$ which is in agreement with the numerical results (see Figure 3.3). Clearly the behaviour calculated in equation (3.13) will not be seen on a lattice smaller than 10^4 (corresponding to $\hat{\rho}_{min} > \frac{1}{10}$) as it comes from small instanton effects.

In their computer simulations, Ishikawa et al. (1983) measured the topological charge, q_L, contained in a box (either the entire 8^4 lattice or a 4^4 sublattice embedded in the 8^4 host lattice)

$$q_L = \sum_{box} \tilde{F} \cdot F$$

for a number of Monte Carlo equilibrium configurations and plotted the resultant histograms for observed values of q_L against frequency. We can show that this frequency histogram corresponds to



<u>Figure 3.3</u>

Distribution of topological charge near $q = 1$ (solid lines). The broken lines show the effect of imposing a minimum instanton scale size $\rho_{min} = R/4$.

our distribution function $P(q)$ as follows. We have from (3.10) that

$$P(q) = \int \mathcal{D}A_\mu \, e^{-S[A]} \, \delta(q - Q[A])$$

$$= \langle \delta(q - Q[A]) \rangle \quad .$$

The Monte Carlo prescription as described in the prologue allows us to write this as the sum over M equilibrium configurations A_n, $n = 1, \ldots, M$

$$P(q) = \frac{1}{M} \sum_{n=1}^{M} \delta(q - Q[A_n]) \tag{3.14}$$

Individual values of Q cannot be measured exactly so we group together results into bins of width δq and plot a histogram. This corresponds to averaging (3.14) over the width of the bin, giving us the frequency $F(q)$ associated with the bin $[q - \frac{\delta q}{2}, \; q + \frac{\delta q}{2}]$

$$F(q)\delta q = \int_{q - \frac{\delta q}{2}}^{q + \frac{\delta q}{2}} P(q')dq'$$

$$= \frac{1}{M} \sum_{n=1}^{M} \int_{q - \frac{\delta q}{2}}^{q + \frac{\delta q}{2}} \delta(q' - Q[A_n])dq'$$

$$= \frac{1}{M} \sum_{n=1}^{M} \Delta(Q[A_n], q)$$

where $\Delta(Q[A_n],q) = \begin{cases} 1 & \text{if } Q[A_n] \in \left[q - \frac{\delta q}{2}, \; q + \frac{\delta q}{2}\right] \\ 0 & \text{otherwise} \end{cases}$

Hence we have

$$F(q)\delta q = \frac{m}{M}$$

where m is the number of measurements of q falling within the bin and M is the number of configurations.

We cannot make any quantitative comparisons with these Monte Carlo results as the data in the range $\frac{1}{2} \lesssim q \lesssim 1$ are too sparse - however, some qualitative remarks can be made. In their search for instantons on the lattice Ishikawa et al. (1983) expected to see a distribution consisting of peaks about the integers (Figure 1 of Ishikawa et al.). In fact they observed distributions which decreased monotonically away from the origin and interpreted this as an indication of the absence of instantons on their lattices. However, the distribution $P(q)$ which we have calculated has the same qualitative behaviour as the measured distributions.

Ishikawa et al. also expected to observe the same qualitative behaviour for SU(2) and SU(3). We have seen, though, that the situation is more complex and on a sufficiently fine lattice it should be possible to observe a marked difference between the two.

3.4 Conclusions and Further Remarks

In this chapter we have calculated the probability density function for the topological charge q contained in a sphere of finite radius for q in the range $\frac{1}{2} \lesssim q \lesssim 1$. Although we worked within the dilute gas approximation there was no need to impose an arbitrary maximum on instanton scale size. We also found our results to be in qualitative agreement with Monte Carlo simulations.

The calculation may be extended to values of $q < \frac{1}{2}$ by allowing instantons with scale sizes greater than R. Again, there will be no need for an arbitrary cutoff as the maximum instanton scale size is still controlled by q and R. Contributions to $q > 1$ come from multi-instanton effects.

There is one feature of the results in this chapter which should be largely unaffected by the lattice and could be tested readily in Monte Carlo simulations. In equation (3.11) we saw that the sector of $P(q)$ dominated by one instanton should scale with R like $R^{\frac{11N}{3}}$. (The effect of the lattice is to introduce a minimum scale size for the instantons which alters the upper limit of the integral in (3.12), as illustrated in the last section. This should not affect the scaling behaviour.) This scaling is already consistent with the Monte Carlo data of Ishikawa et al. (1983) in which the topological charge distribution for SU(2) was measured on both the full 8^4 lattice and also a 4^4 sub-lattice embedded in the 8^4 host lattice for the same value of the lattice coupling. (The advantage of using an embedded sub-lattice is that in this way one can reduce the systematic errors due to finite size effects.) For the region in which single instantons dominate we would expect a scaling by a factor of $2^{22/3} \simeq 160$ between the two distributions and this is consistent with the measured data.

Finally, the results of this chapter may also be used to test the picture of the QCD vacuum developed by Callan, Dashen and Gross (see Callan et al., 1979 and Callan, 1980). This proposes that the length scale of the lowest lying hadrons is sufficiently small for the semi-classical approximations - and in particular for the dilute gas approximation - still to be valid, with the (incalculable) strong coupling physics setting in at larger length scales. If we were to measure the topological charge distribution at successively larger values of the lattice coupling (and hence with successively larger values of R) we should observe first the breakdown of single instanton dominance and then the breakdown of the dilute gas approximation and the onset of strong coupling physics. Comparison of the length scale at which this occurs with the length scale of the lowest lying hadrons (from Monte Carlo hadron simulations) should reveal whether or not the semi-classical approximation is still valid at these length scales and so indicate whether or not it can ever describe any hadron physics.

PART II

GAUGE THEORIES ON A SPACE TIME LATTICE

PART II

CHAPTER 4

INTRODUCTION TO LATTICE GAUGE THEORY

4.1 General Introduction

Much recent interest in physics has been in the construction and investigation of suitable lattice models of the gauge theories believed to describe the fundamental forces of nature. Some of the attractions of a lattice regularization have been outlined in the Prologue but there are various difficulties which have to be faced. Of these, some of the most serious are the problems encountered in the formulation of these lattice theories at a fundamental level. It is by no means clear how to implement a lattice regularization for general theories and we must adopt a partially constructive approach from basic principles.

Imagine a lattice which is a triangulation of the manifold on which the theory is defined. This manifold we will take to be Euclidean in signature (having already done a Wick rotation from a Minkowski-type space as described in the prologue). This lattice we shall take to be hypercubic with a lattice spacing a, as this is conceptually the easiest with which to work and to formulate theories. (Technically it should not matter what triangulation is used.) The path integral is replaced by a countable product of integrals at each lattice site and the integral over the manifold is replaced by a sum at each site

$$\int \mathcal{D}\phi \; e^{-\int d^4x \mathcal{L}^c[\phi(x)]} \quad \rightarrow \quad \int \left[\prod_n d\phi(n) \right] e^{-\sum_n \mathcal{L}^L[\phi(n)]} \quad . \tag{4.1}$$

The problem with which we shall be concerned in this chapter and the next is the finding of suitable lattice analogues $\mathcal{L}^L$ of the continuum action density $\mathcal{L}^c$. The lattice site referred to by n means the point with co-ordinates a (n_1, n_2, n_3, n_4) and we shall use $n{+}\mu$ to label the site $\underline{n} + a.\underline{e}_\mu$. (For a general overview of lattice gauge theory see Wilson (1974), Kogut (1979) and Kogut (1983).)

We can formulate a free scalar field theory on the lattice with comparative ease. Consider the action

$$- \sum_{n,\mu} \left[\phi^\dagger(n)[\phi(n+\mu) + \phi(n-\mu)] \right] + \sum_n m_o^2 \; \phi^\dagger(n)\phi(n) \tag{4.2}$$

where ϕ is some N-component field. Then as we take the continuum limit, $\sum_n \rightarrow \int \frac{d^4x}{a^4}$; and it is necessary also to rescale the field ϕ (which in the continuum has dimension of 1 mass unit)

$$\phi(n) \quad \rightarrow \quad \frac{a}{\sqrt{2}} \, \phi(x) \quad .$$

Expanding $\phi(n{+}\mu)$ as we take the limit $a \rightarrow 0$ we obtain

$$\begin{aligned} &\tfrac{1}{2}\int d^4x \left[-\phi^\dagger(x)\nabla^2\phi(x) + m^2 \; \phi^\dagger(x)\phi(x) \right] \\ &= \int d^4x \left[\tfrac{1}{2}[\partial_\mu\phi]^\dagger[\partial_\mu\phi] + \tfrac{1}{2}m^2\phi^\dagger(x)\phi(x) \right] \end{aligned} \tag{4.3}$$

where $m^2 = \dfrac{m_o^2 - 8}{a^2}$ which has the units of mass squared.

Hence the continuum limit of (4.2) is the free scalar field theory. Now we can re-write the action (4.2) as

$$- \sum_{n,\mu} \left[(\phi^{\dagger}(n)\phi(n+\mu) - 2\phi^{\dagger}(n)\phi(n) + \phi^{\dagger}(n)\phi(n-\mu) \right] + \sum_{n} m^2 \phi^{\dagger}(n)\phi(n)$$

(where now $m^2 = m_o^2 - 8$) and by taking the Fourier transform we arrive at the propagator $S(p)$

$$\begin{aligned} S(p)^{-1} &= - \sum_{\mu} (e^{ip_\mu a} - 2 + e^{-ip_\mu a}) + m^2 \\ &= 4 \sum_{\mu} \sin^2 \frac{p_\mu a}{2} + m^2 \\ &\rightarrow p^2 a^2 + m^2 \quad \text{as} \quad a \rightarrow 0 \end{aligned}$$

which gives us the correct propagator in the limit $a \rightarrow 0$, with a particle of mass m/a. Now p_μ is the momentum which must exist in the Brillouin zone $-\frac{\pi}{a} < p_\mu \leq \frac{\pi}{a}$ and so we see there is just one lowest lying state (corresponding to just one particle state) at $p_\mu = 0$, all μ.

(4.2) clearly has a global U(N) symmetry (O(N) if the fields are restricted to be real). Let $\Omega \in$ symmetry group: then the action (4.2) is invariant under the transformation

$$\phi(n) \rightarrow \Omega\, \phi(n) \quad .$$

We can make this global symmetry into a local one, and so construct a gauge theory by allowing Ω to vary from site to site

$$\phi(n) \rightarrow \Omega(n)\, \phi(n). \tag{4.4}$$

Then the mass term in (4.2) is invariant but the nearest-neighbour coupling is not: we must introduce a field $U_\mu(n)$ which transforms as

$$U_\mu(n) \to \Omega(n)\, U_\mu(n)\, \Omega^{-1}(n+\mu) \qquad (4.5)$$

and then both terms of

$$\phi^\dagger(n)U_\mu(n)\phi(n+\mu) \quad + \quad \phi^\dagger(n+\mu)U_\mu^\dagger(n)\phi(n)$$

will be invariant. $\{U_\mu(n)\}$ then become the dynamical variables of the gauge field; from the transformation properties (4.5) we see that they may be taken as the group elements themselves, rather than elements of the Lie Algebra, so we may write $U_\mu(n) = \exp[ig\, A_\mu(n)]$. These variables may also be thought of as being defined on the links of the lattice ($U_\mu(n)$ on the link between n and $n+\mu$, etc.). It is not so easy to write down the lattice analogue for the kinetic term for the gauge variables: however it turns out that the simplest gauge invariant operator possible does indeed have the right continuum limit: this is the product over an elementary plaquette of the link variables

$$\sum_{\substack{\mu,\nu \\ \mu\neq\nu}} \mathrm{Tr}\left[U_\mu(n)U_\nu(n+\mu)U_\mu^\dagger(n+\nu)U_\nu^\dagger(n) \; + \; \text{h.c.}\right] \qquad (4.6)$$

This is clearly gauge invariant, and it is not difficult to show (see, for example, Wilson (1974)) that the continuum limit is the standard gauge kinetic term

$$\mathrm{Tr}\; F_{\mu\nu}\, F^{\mu\nu} \quad .$$

Adding (4.6) to the action (4.2) we have a model to describe scalar gauge theory. We are now dealing with an interacting theory so the true continuum limit (in which we remove the regularization $(a \to 0)$ after renormalization has taken place) is non-trivial: however we can still take a naive limit by letting $a \to 0$ in which case we recover the bare and unregularized continuum theory.

We have only presented here the general features of lattice gauge theories: the matter fields are defined on the sites and the gauge fields on the links (with appropriate gauge transformation properties) and there is no difficulty with the kinetic terms for the gauge and the scalar fields. For further details and in particular for numerical techniques used in simulations see Kogut (1979) and Kogut (1983). We now turn our attention to the difficulties associated with formulating lattice theories with fermions.

4.2 Lattice Fermions and Species Doubling

A free fermion with mass m has the continuum action (in Euclidean space)

$$- \int d^4x \left[\bar{\psi}(x)\gamma_\mu \partial_\mu \psi(x) + m \bar{\psi}(x)\psi(x) \right] . \tag{4.7}$$

To discretize this, we replace the derivative with differences. Now the standard prescription (see, for example, Rabin, 1982) is

$$\begin{aligned} &\text{either} \quad \partial_\mu \psi(n) \to \frac{\psi(n+\mu) - \psi(n)}{a} \\ &\text{or} \quad \partial_\mu \psi(n) \to \frac{\psi(n) - \psi(n-\mu)}{a} \end{aligned} \tag{4.8}$$

and in this way we can derive the Laplacian, (using both)

$$\partial_\mu \partial^\mu \psi(n) \rightarrow \frac{\psi(n+\mu) - 2\psi(n) - \psi(n-\mu)}{a^2}$$

confirming the relationship between (4.2) and (4.3). However, we have only one derivative in (4.7) and so cannot use the asymmetric prescriptions above, but rather a symmetric combination of them

$$\partial_\mu \psi(n) \rightarrow \frac{\psi(n+\mu) - \psi(n-\mu)}{2a} \quad . \tag{4.9}$$

Upon rescaling the fields by $\psi(n) = (\frac{a^3}{2K})^{\frac{1}{2}} \psi(x)$ (to make a dimensionless field for the lattice action) we have the discretized version

$$- K \sum_{n,\mu} \left[\overline{\psi}(n)\gamma_\mu(\psi(n+\mu) - \psi(n-\mu))\right] - \sum_n \overline{\psi}(n)\psi(n) \tag{4.10}$$

where the identification $m = \frac{1}{2aK}$ has been made. (4.10) may look satisfactory, but it conceals a major difficulty which will be exposed when we examine the propagator. Taking a Fourier transform as before, the propagator $S(p)$ is

$$\begin{aligned} S(p)^{-1} &= K \sum_\mu \left[(e^{ip_\mu a} - e^{-ip_\mu a})\gamma_\mu\right] + 1 \\ &= 2iK \sum_\mu \gamma_\mu \sin p_\mu a + 1 \qquad (4.11) \\ &\rightarrow 2iKap_\mu\gamma^\mu + 1 \quad \text{as} \quad a \rightarrow 0 \end{aligned}$$

or, with rescaling, $ip_\mu\gamma^\mu + m$ with $m = \frac{1}{2aK}$. Thus the propagator has the right continuum limit, but since $-\frac{\pi}{a} < p_\mu \leqslant \frac{\pi}{a}$

we see there is a lowest lying state $(\sin p_\mu a = 0)$ not only for $p_\mu = 0$, all μ, but also if any of the four components of p_μ take the value $\frac{\pi}{a}$. Thus there are sixteen such poles in the propagator and the action (4.10) describes not one but sixteen fermionic degrees of freedom. This is the notorious fermion doubling problem and we shall return to it shortly. First, however, we should notice that (4.10) can be turned into a gauge theory in exactly the same way as can (4.2), yielding the action

$$- K \sum_{n,\mu} \left[\bar{\psi}(n)U_\mu(n)\gamma_\mu\psi(n+\mu) - \bar{\psi}(n+\mu)U_\mu^\dagger(n)\gamma_\mu\psi(n)\right] - \sum_n \bar{\psi}(n)\psi(n). \quad (4.12)$$

(4.12) is the standard "naive" lattice action for a fermion, describing sixteen degenerate states in the continuum limit.

4.3 Resolution of the Doubling Problem

There are two main techniques which are used to resolve this multiple degeneracy. The simplest one was introduced by Wilson (Wilson, 1974) and involves introducing a momentum dependent mass term

$$Kr \sum_{n,\mu} \left[\bar{\psi}(n)U_\mu(n)\psi(n+\mu) + \bar{\psi}(n+\mu)U_\mu^\dagger(n)\psi(n)\right] \quad (4.13)$$

which gives the unwanted states at the edge of the Brillouin zone a mass of the order of the cutoff. The propagator becomes

$$S(p)^{-1} = 2iK \sum_\mu \gamma_\mu \sin p_\mu a + 1 - 2Kr \sum_\mu \cos p_\mu a \quad .$$

The state at $p_\mu = 0$ now has mass

$$m_o = \frac{1 - 8Kr}{2aK}$$

but those with one or more momenta component at the edge of the Brillouin zone pick up a higher mass

$$m = m_o + \frac{2r}{a} \cdot k$$

for $k = 1$ (four states), 2 (six states), 3 (four states) and 4 (one state). Thus the extra states have masses boosted to the order of the cutoff and so should not influence the low energy physics.

The disadvantage of this approach is that the Wilson "r-term" (4.13) explicitly breaks chiral symmetry. As we take the naive continuum limit this term vanishes (it is $O(a)$ as $a \to 0$) but this is no guarantee that chiral symmetry will be restored in the non-trivial (interacting) theory. Furthermore, to ensure that the continuum has finite mass particles, there will have to be fine tuning of the "hopping parameter" K (see Wilson, 1974 and Kogut, 1983).

The other main technique is the one put forward by Kogut and Susskind (see, for example, Kogut 1983). Here the Dirac operator is diagonalized (in spinor space) by a special transformation (Kawamoto and Smit, 1981)

$$\psi(n) = \gamma_o^{|n_o|} \gamma_1^{|n_1|} \gamma_2^{|n_2|} \gamma_3^{|n_3|} \chi(n) .$$

The action for χ is then identical for each spin component and so the degeneracy has been reduced from 16 to 4. (A deeper understanding of this spin-diagonalization is presented later.) We shall not investigate this formulation any further as we shall be interested mainly in Wilson's prescription: further details are available in the articles already referenced.

4.4 Deeper Understanding of Fermion Doubling

The fermion doubling problem is now understood at a deep level by consideration of the Dirac-Kähler equation. This is a generalization of the Dirac equation in the sense that it is not only a square root of the Laplacian, as is the Dirac equation, but it is also the full realization of the Clifford algebra of which the γ matrices form a representation. The fields are defined in terms of the differential forms or, alternatively, 4×4 matrices - that is, sixteen component objects. (For a review of the Dirac-Kähler equation, both in the continuum and the lattice, in these two formulations, see Becher and Joos (1982) and Rabin (1982) respectively.) The equation can then be written very simply and it turns out that in the continuum it can be split into four independent copies of the Dirac equation, each describing a four-component spinor. There is no degeneracy introduced by going to the lattice but these four copies no longer decouple. Consequently the 16-fold doubling is made manifest.

The approach also gives insight into why the Kawamoto-Smit spin diagonalization is possible. Consider the group, M, of transformations mapping the degenerate particle states into each other: clearly this commutes with the Dirac operator. However, it turns out to be another representation of the Clifford algebra, as are the γ matrices. Since all such representations are equivalent, there must be a similarity transformation T

$$M \xrightarrow{T} \{\gamma\} .$$

In the basis obtained by this transformation, the Dirac operator will commute with a representation of the γ matrices (that is, M). However, the γ matrices generate the space of all 4×4 matrices: hence the Dirac operator must be a multiple of the identity in this basis (Becher and Joos, 1982).

There is also a deep connection between chiral symmetry and fermion doubling. It can be established that doubling cannot be avoided if the chiral symmetry is to be preserved on the lattice (Wilson fermions escape it by explicitly breaking the chiral symmetry). This connection can be shown by studying quantum effects (anomalies) - see Karsten and Smit (1981) and also by geometrical arguments at a classical level (Rabin, 1982).

This link between fermion doubling and chiral symmetry lies at the heart of the no-go theorems of Nielsen and Ninomiya (1981(a) and (b)) which present an obstruction to putting neutrinos on the lattice. It is this difficulty to which we turn out attention in the following chapter.

CHAPTER 5

LATTICE ACTIONS WITH LEFT-RIGHT ASYMMETRY

5.1 The Neutrino Problem

The impossibility of putting neutrinos onto the lattice was established by Nielsen and Ninomiya in 1981. The obstruction follows from the fermion doubling described earlier. If one tries to formulate naively an action for purely left-handed massless fermion fields

$$S_F = - \sum_{n,\mu} [\bar{L}(n)U_\mu(n)\gamma_\mu L(n+\mu) - \bar{L}(n+\mu)U_\mu^\dagger(n)\gamma_\mu L(n)], \qquad (5.1)$$

then the no-go theorems (Nielsen and Ninomiya, 1981(a) and (b)) require that the states which appear through doubling yield a final spectrum in which there are equally many left and right handed states. Consequently, the conclusion drawn is that the future looks bleak for lattice descriptions of weak interactions, involving, as they do, purely left handed neutrinos and other fermions with their left and right handed parts transforming according to different representations of the gauge group.

This chapter introduces a class of theories which circumvents these difficulties by requiring that only the currents be left-handed, rather than the elementary particles. The doubling problem is surmounted by making use of Wilson's prescription as described in Chapter 4 - then the doubled states do appear but they have masses of the order of the cutoff ($^1/a$) and so are not expected to contribute to physical processes.

Consider the Wilson action

$$S_F = -\sum_n \bar{\psi}(n)\psi(n) - K\sum_{n,\mu}\Big[\bar{\psi}(n)\gamma_\mu U_\mu(n)\psi(n+\mu) - \bar{\psi}(n+\mu)\gamma_\mu U_\mu^\dagger(n)\psi(n)\Big]$$

$$+ Kr\sum_{n,\mu}\Big[\bar{\psi}(n)U_\mu(n)\psi(n+\mu) + \bar{\psi}(n+\mu)U_\mu^\dagger(n)\psi(n)\Big]. \qquad (5.2)$$

Then we can separate the left and the right handed parts by writing
$L(n) = (\frac{1-\gamma_5}{2})\psi(n)$, $R(n) = (\frac{1+\gamma_5}{2})\psi(n)$ so that $L(n)+R(n) = \psi(n)$.

Recalling the properties of the left and right projection matrices

$$(\frac{1\pm\gamma_5}{2})^2 = (\frac{1\pm\gamma_5}{2}) \quad ,$$

$$(\frac{1+\gamma_5}{2})(\frac{1-\gamma_5}{2}) = 0 \quad ,$$

$$(\frac{1\pm\gamma_5}{2})\gamma_\mu = \gamma_\mu(\frac{1\mp\gamma_5}{2})$$

the action becomes

$$S_F = -K\sum_{n,\mu}\Big[\bar{L}(n)\gamma_\mu U_\mu(n)L(n+\mu) - \bar{L}(n+\mu)\gamma_\mu U_\mu^\dagger(n)L(n)\Big]$$

$$-K\sum_{n,\mu}\Big[\bar{R}(n)\gamma_\mu U_\mu(n)R(n+\mu) - \bar{R}(n+\mu)\gamma_\mu U_\mu^\dagger(n)R(n)\Big]$$

$$-\sum_n\Big[\bar{L}(n)R(n) + \bar{R}(n)L(n)\Big]$$

$$+ Kr\sum_{n,\mu}\Big[\bar{L}(n)U_\mu(n)R(n+\mu) + \bar{R}(n)U_\mu(n)L(n+\mu)$$
$$+ \bar{R}(n+\mu)U_\mu^\dagger(n)L(n) + \bar{L}(n+\mu)U_\mu^\dagger(n)R(n)\Big] \qquad (5.3)$$

On looking at (5.3) for the first time one's initial reaction might be that it describes two fermions, L and R, of the form (5.1) with a number of interactions between them. This is a rather misleading approach as it suggests that both L and R undergo doubling, lose their chirality and interact with each other through the interaction terms. However, we know that (5.3) can be re-assembled to form (5.2) and that the interaction terms are then seen to be momentum dependent mass terms, lifting the degeneracy of the doubled states by giving them masses of the order of the cutoff. If we ignore these super-massive states, then we see that the spectrum of (5.3) is a single fermion, $\psi = L+R$ with mass $m = \frac{1-8Kr}{2aK}$ and with the L and the R fields being genuinely the left and the right handed parts of the field ψ. The conclusion to be drawn here is that care has to be taken in examining the effect of the interaction terms before it is possible to say what the spectrum is.

In the models to be presented later in this chapter the left and the right handed parts of the field transform according to different representations of the gauge group. Then in a phase in which the gauge symmetry is spontaneously broken, the interaction terms become a momentum dependent mass term and we are able to construct a Wilson fermion so that the doubled states will not concern us. We are left with gauge interactions through currents which may have a specific handedness as the fields concerned (L and R) are the left and right parts of the undoubled field ψ.

5.2 A Left-Right Asymmetric Action

Consider a model with a gauge group G and fermionic fields L and R such that

$$L = \left(\frac{1-\gamma_5}{2}\right)L, \qquad R = \left(\frac{1+\gamma_5}{2}\right)R$$

i.e. L and R are purely left and right handed. Then let L transform according to some representation $U_\mu(n)$ of the gauge group G and for simplicity let R transform trivially. Then under a gauge transformation $\Omega(n) \in$ {the representation of G}

$$\begin{aligned} L(n) &\to \Omega(n)\, L(n) \\ \bar{L}(n) &\to \bar{L}(n)\, \Omega^{-1}(n) \\ U_\mu(n) &\to \Omega(u) U_\mu(n) \Omega^{-1}(n+\mu) \\ R(n) &\to R(n) \quad . \end{aligned}$$

Our first naive attempt at the fermionic part of the lattice action might be

$$\begin{aligned} S_F = & -K \sum_{n,\mu} \left[\bar{L}(n)\gamma_\mu U_\mu(n) L(n+\mu) - \bar{L}(n+\mu)\gamma_\mu U_\mu^\dagger(n) L(n)\right] \\ & -K \sum_{n,\mu} \left[\bar{R}(n)\gamma_\mu R(n+\mu) - \bar{R}(n+\mu)\gamma_\mu R(n)\right] \end{aligned} \tag{5.4}$$

However, as described in the last section, both L and R will undergo doubling as in (5.1) and we are left with 16 fermions (8 left and 8 right) transforming under the representation and 16 fermions (8 left and 8 right) transforming trivially. We need to introduce mass terms and in particular the momentum dependent Wilson term. However, mass terms are like $\bar{L}\,R$ and this is not gauge invariant, so we are led to introduce a scalar ϕ, transforming in the same way under G as L, so that we can write down the interaction terms

$$
\begin{aligned}
& - \sum_n \left[\bar{L}(n)\phi(n)R(n) + \bar{R}(n)\phi^\dagger(n)L(n)\right] \\
& + Kr \sum_{n,\mu} \Big[\bar{L}(n)\phi(n)R(n+\mu) + \bar{R}(n)\phi^\dagger(n+\mu)L(n+\mu) \\
& \qquad + \bar{R}(n+\mu)\phi^\dagger(n)L(n) + \bar{L}(n+\mu)\phi(n+\mu)R(n)\Big]
\end{aligned}
\tag{5.5}
$$

which are gauge invariant. Putting these together with (5.4) we must now consider the full action

$$
\begin{aligned}
S = & -K \sum_{n,\mu} \left[\bar{L}(n)\gamma_\mu U_\mu(n)L(n+\mu) - \bar{L}(n+\mu)\gamma_\mu U_\mu^+(n)L(n)\right] \\
& -K \sum_{n,\mu} \left[\bar{R}(n)\gamma_\mu R(n+\mu) - \bar{R}(n+\mu)\gamma_\mu R(n)\right] \\
& - \sum_n \left[\bar{L}(n)\phi(n)R(n) + \bar{R}(n)\phi^\dagger(n)L(n)\right] \\
& + Kr \sum_{n,\mu} \left[\bar{L}(n)\phi(n)R(n+\mu) + R(n)\phi^\dagger(n+\mu)L(n+\mu) + h.c.\right] \\
& + \beta_H \sum_{n,\mu} \left[\phi^\dagger(n)U_\mu(n)\phi(n+\mu) + \phi^\dagger(n+\mu)U_\mu^\dagger(n)\phi(n)\right] \\
& + \beta \operatorname{Re} \operatorname{Tr} \sum_n \{ \prod_{plaq} U_\mu(n)\}.
\end{aligned}
\tag{5.6}
$$

We have included in (5.6) suitable kinetic terms for the scalar field ϕ and for the gauge fields. However, as demonstrated in the last section we cannot comment on the spectrum or the nature of the doubling: the effect of the interaction terms (5.5) must be examined first. Before we can do that, however, we must turn our attention to a simpler model than (5.6) - the Abelian Higgs model.

5.3 Abelian Higgs Models

Consider a model with an Abelian gauge group G and with scalar fields ϕ transforming according to some representation $U_\mu(n)$ of the group. The action for this is

$$S = \beta \operatorname{Re} \operatorname{Tr} \sum_n \left[\prod_{\text{plaq}} U_\mu(n) \right]$$

$$+ \beta_H \sum_{n,\mu} \left[\phi^\dagger(n) U_\mu(n) \phi(n+\mu) + \phi^\dagger(n+\mu) U_\mu^\dagger(n) \phi(n) \right] . \qquad (5.7)$$

The phase diagram of this model has been studied in detail by Fradkin and Shenker (1979) for discrete groups G such as Z_n as well as for U(1). The limiting cases of the model are as follows:-

(a) $\beta = \infty$: here the gauge fields will be frozen out to pure gauge configurations. Picking an axial gauge in which $U_\mu(n) = I$ the action reduces to

$$S = \beta_H \sum_{n,\mu} \left[\phi^\dagger(n) . \phi(n+\mu) + \phi^\dagger(n+\mu) . \phi(n) \right] .$$

For $G = U(1)$ this is the X-Y model - there is a global U(1) invariance and there is a phase transition at $\beta_H = \beta_H^c$ (see R. Griffiths (1972) and J. José et al. (1977)). For $\beta_H > \beta_H^c$ the global U(1) is spontaneously broken and the Higgs field develops a vacuum expectation value.

(b) $\beta_H = 0$: here we are left with the pure gauge theory,

$$S = \text{Re Tr} \sum_n \left[\prod_{\text{plaq}} U_\mu(n) \right]$$

as the Higgs fields decouple. This too has a phase transition at $\beta = \beta_c$: this is the deconfining phase transition between strong coupling ($\beta < \beta_c$) and non-confining Coulomb type weak coupling ($\beta > \beta_c$). (See Kogut and Susskind, (1975) and J. Jose et al. (1977)).

The work of Fradkin and Shenker was to investigate the interior of the phase diagram. They found that it is divided into three regimes (see Figs. 5.1, 5.2):

(i) Higgs mechanism regime, characterized by massive gauge bosons and no confinement of gauge charge (β_H, β both large: $\beta_H > \beta_H^c$, $\beta > \beta_c$).

(ii) Coulomb phase: the gauge bosons are massless and there is no confinement, giving a long-range Coulomb force between static sources (β_H small β large: $\beta_H < \beta_H^c$, $\beta > \beta_c$).

(iii) Confinement regime, characterized by confinement of gauge charge - the spectrum consists of gauge singlets (β small: $\beta < \beta_c$).

The Coulomb phase is always separated from the other two regimes by phase transitions, but the Higgs and confining regimes need not be separate phases. It turns out that if the Higgs fields transform in the fundamental representation, then there is no phase transition between them - they are two regimes with differing spectra within the same phase (see Fig. 5.1). Alternatively, if the Higgs fields transform in a higher representation we have the case shown in Fig. 5.2 where the Higgs and confining regimes are

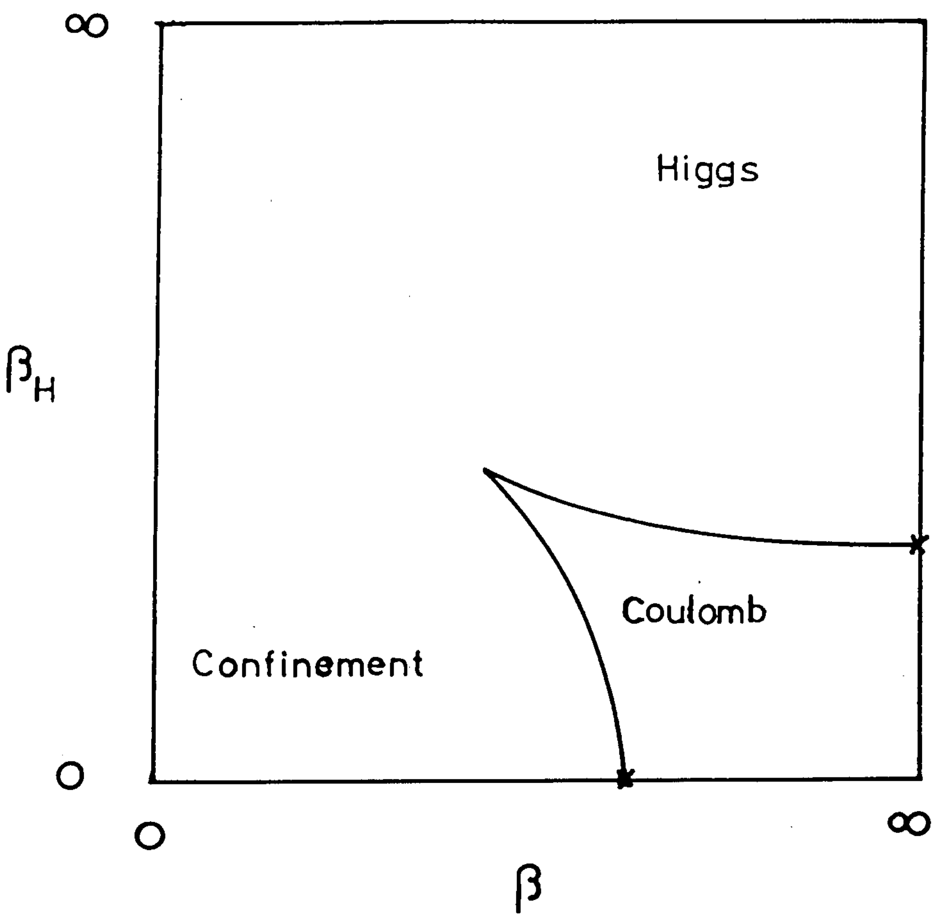


Figure 5.1

Phase diagram for the Abelian Higgs model with the Higgs fields in the fundamental representation. The solid lines are phase transitions:- there is no phase transition between the Higgs and confining regimes.

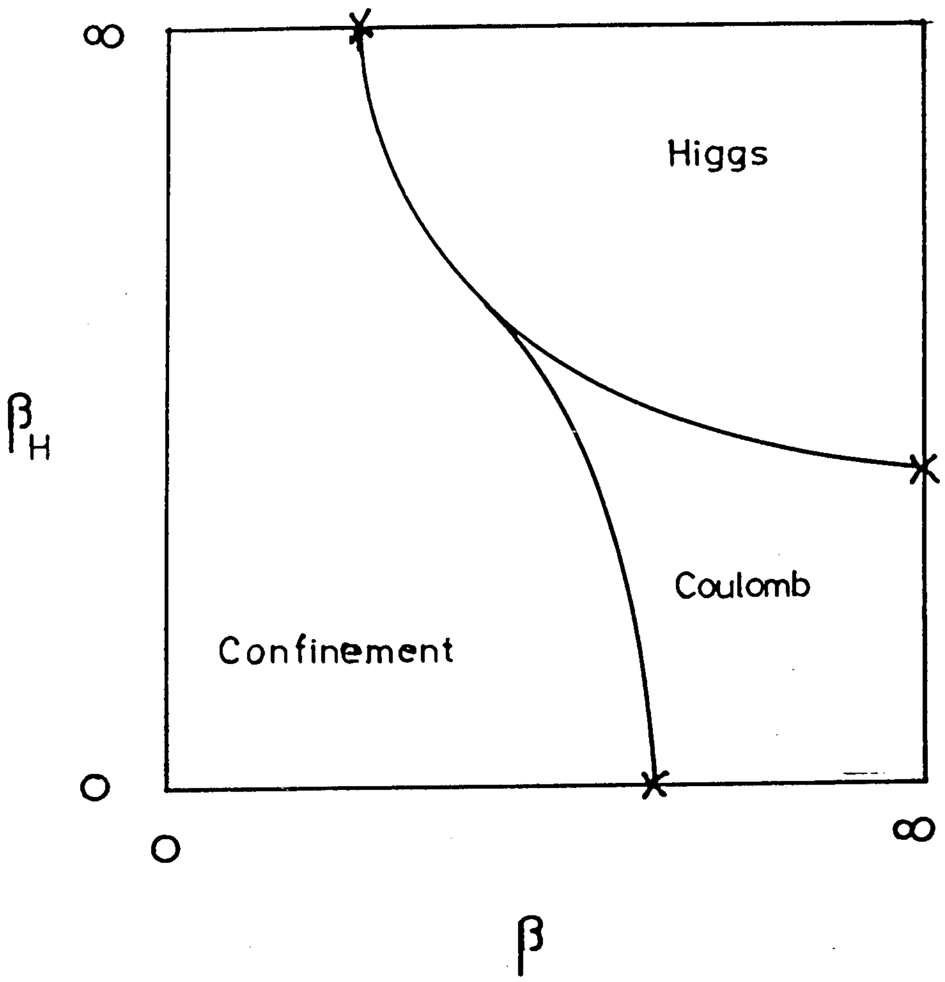


Figure 5.2

Phase diagram for the Abelian Higgs model with the Higgs fields in the adjoint representation. The Higgs and confining regimes are separated by phase transitions.

separated by a phase transition.

These results are well supported by Monte Carlo studies undertaken by a number of researchers (see Creutz (1980) and Jongeward et al. (1980) for Z_n groups and Bowler et al. (1981) for $U(1)$).

It turns out (Fradkin and Shenker, (1979)) that these results persist if the gauge group is no longer chosen to be Abelian, such as the $SU(N)$ groups, at any rate if the dimensionality of space-time is high enough $(d > 4)$. The difference for $d = 4$ is that, if the Higgs fields transform in the fundamental representation, the Coulomb phase is no longer present.

5.4 Models with Left-Right Asymmetry: Analysis of Spectra

We now return to (5.6) armed with the knowledge of Abelian Higgs models reviewed in the last section. The model we are interested in is more complicated because it includes the fermions: nevertheless the general features will be reproduced and in particular the existence of a Higgs phase for β and β_H both large. It is the spectrum in this sector that we wish to examine. Large β_H implies that the scalar fields freeze out $(\phi^\dagger\phi = 1)$ so following Fradkin and Shenker we choose a unitary gauge in which $\phi = \phi_o$, say (with $\phi_o^\dagger\phi_o = 1$) and examine the spectrum. We shall do this for a number of examples with differing groups and representations. For a review of the group theory used in this section see D.B. Lichtenberg (1978), and for the continuum Higgs models see J.C. Taylor (1978).

(i) SU(2) with fields in the fundamental representation

Both L and ϕ are complex doublets, R is a singlet and $U_\mu(n) = \exp[iga\, \underset{\sim}{A}_\mu \cdot \underline{\tau}]$ where $\{\tau_i\}$ are the Pauli spin matrices.

G = SU(2) so dimG = 3. The vacuum manifold

$$M_o = \{\phi_o : \phi_o^\dagger \phi_o = 1\}$$

has dim M_o = 3. Now if H is the little group (the residual symmetry group after symmetry breaking), we have $G/H \cong M_o$ and so

$$\dim G - \dim H = \dim M_o \ . \tag{5.8}$$

Hence dim H = 0 and there is no residual symmetry left. Therefore all the gauge bosons will acquire a mass. Let $\phi_o = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and expand $U_\mu(n) = \exp[iga\, \underline{\tau} \cdot \underset{\sim}{A}_\mu(n)]$. The gauge boson mass term is from

$$\beta_H \left[\phi_o^\dagger\, U_\mu(n) \phi_o + \phi_o^\dagger\, U_\mu^\dagger(n) \phi_o \right]$$

$$= 2\beta_H (0,1) \left[1 - g^2 a^2 [\underset{\sim}{A}_\mu(n) \cdot \underline{\tau}]^2 \right] \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ .$$

Now the Pauli matrices obey $\tau^i \tau^j = \delta^{ij}\, \mathbb{1} + i\, \varepsilon^{ijk}\, \tau^k$ so

$$A_\mu^i(n) A_\mu^j(n) \quad \tau^i \tau^j = \underset{\sim}{A}_\mu^2 \quad \mathbb{1}$$

and so the mass term is

$$2g^2 a^2 \beta_H (\underset{\sim}{A}_\mu)^2 \ .$$

That is, all three bosons acquire the same mass.

Now let us examine the fermionic sector. Putting $L = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$ and expanding the action we get

$$\begin{aligned}
S_F = &- K \sum_{n,\mu} \left[\bar{L}_1(n)\gamma_\mu \; L_1(n+\mu) - \bar{L}_1(n+\mu)\gamma_\mu \; L_1(n) \right] \\
&- K \sum_{n,\mu} \left[\bar{L}_2(n)\gamma_\mu \; L_2(n+\mu) - \bar{L}_2(n+\mu)\gamma_\mu \; L_2(n) \right] \\
&- K \sum_{n,\mu} \left[\bar{R}(n)\gamma_\mu \; R(n+\mu) - \bar{R}(n+\mu)\gamma_\mu \; R(n) \right] \\
&- K \sum_{n,\mu} \left[(\bar{L}(n)\gamma_\mu \underline{\tau} \; L(n+\mu) + \bar{L}(n+\mu)\gamma_\mu \; \underline{\tau} \; L(n)) \cdot ig\underline{A}_\mu(n) \right] \\
&+ Kr \sum_{n,\mu} \left[\bar{L}(n) \begin{pmatrix} 0 \\ 1 \end{pmatrix} R(n+\mu) + \bar{R}(n)(0,1) \; L(n+\mu) \right. \\
&\qquad \left. + \bar{L}(n+\mu) \begin{pmatrix} 0 \\ 1 \end{pmatrix} R(n) + \bar{R}(n+\mu)(0,1) \; L(n) \right] \\
&\qquad - \sum_{n} \left[\bar{L}(n) \begin{pmatrix} 0 \\ 1 \end{pmatrix} R(n) + \bar{R}(n)(0,1) \; L(n) \right]
\end{aligned}$$

$$\begin{aligned}
= &- K \sum_{n,\mu} \left[\bar{L}_2(n)\gamma_\mu \; L_2(n+\mu) - \bar{L}_2(n+\mu)\gamma_\mu \; L_2(n) \right] \\
&- K \sum_{n,\mu} \left[\bar{R}(n)\gamma_\mu \; R(n+\mu) - \bar{R}(n+\mu)\gamma_\mu \; R(n) \right] \\
&- \sum_{n} \left[\bar{L}_2(n)R(n) + \bar{R}(n)L_2(n) \right] \\
&+ Kr \sum_{n,\mu} \left[\bar{L}_2(n)R(n+\mu) + \bar{R}(n)L_2(n+\mu) \right. \\
&\qquad \left. + \bar{R}(n+\mu)L_2(n) + \bar{L}_2(n+\mu)R(n) \right] \\
&- K \sum_{n,\mu} \left[\bar{L}_1(n)\gamma_\mu L_1(n+\mu) - \bar{L}_1(n+\mu)\gamma_\mu L_1(n) \right] \\
&- igK \sum_{n,\mu} \left[\bar{L}(n)\gamma_\mu \; \underline{\tau} L(n+\mu) + \bar{L}(n+\mu)\gamma_\mu \; \underline{\tau} \; L(n) \right] \cdot \underline{A}_\mu(n)
\end{aligned} \tag{5.9}$$

Comparison with (5.3) reveals what we have: if we put $\psi = L_2+R$ we can construct a Wilson fermion: the action becomes

$$
\begin{aligned}
S_F = & -\sum_n \bar{\psi}(n)\psi(n) + K \sum_{n,\mu}\left[\bar{\psi}(n)(r-\gamma_\mu)\psi(n+\mu) + \bar{\psi}(n+\mu)(r+\gamma_\mu)\psi(n)\right] \\
& - K \sum_{n,\mu}\left[\bar{L}_1(n)\gamma_\mu L_1(n+\mu) - \bar{L}_1(n+\mu)\gamma_\mu L_1(n)\right] \\
& - K\, ig \sum_{n,\mu} \underline{j}_\mu(n)\cdot\underline{A}_\mu(n) \qquad (5.10)
\end{aligned}
$$

where $\underline{j}_\mu(n)$ is the current

$$
\underline{j}_\mu(n) = \left[\bar{L}(n)\gamma_\mu \underline{\tau}\, L(n+\mu) + \bar{L}(n+\mu)\gamma_\mu \underline{\tau}\, L(n)\right] .
$$

The spectrum of (5.10) is as follows:-

(a) a Wilson fermion $\psi = L_2+R$, mass $= \frac{1-8Kr}{2aK}$

(b) a massless naive Weyl fermion L_1

(c) a coupling to the gauge bosons $\underline{A}_\mu$ through a current $\underline{j}_\mu$.

(d) There are also the three massive gauge bosons $\underline{A}_\mu$.

Now the Wilson fermion, $\underline{\psi}$, will undergo doubling but the partners so produced will have a mass of the order of the cutoff and so need not be considered in the current coupling to the gauge bosons. The other fermion, L_1, will undergo ordinary doubling however and so will lose its chirality, as described in Section 5.1. Consequently, the current $\underline{j}_\mu$ will not be genuinely left-handed. The parts of it involving L_2, which is the left-handed part of ψ, will be a true left-handed current, but the part coupling L_1 to itself will be mixed - both left and right-handed parts - as L_1

is not specifically left-handed.

Before we examine some other example models, a comment can be made about the "Wilson term" coupling L to R and the Higgs field. The interaction

$$\bar{L}(n)\phi(n)R(n+\mu) + \bar{R}(n)\phi^{\dagger}(n+\mu)L(n+\mu) + \text{h.c.} \tag{5.11}$$

was used (see (5.5)), whereas

$$\bar{L}(n)U_{\mu}(n)\phi(n+\mu)R(n+\mu) + \bar{R}(n)\phi^{\dagger}(n)U_{\mu}(n)L(n+\mu) + \text{h.c.} \tag{5.12}$$

might just as well have been used.

Both (5.11) and (5.12) have the same continuum limit. The lowest order term $\bar{L}(x)\phi(x)R(x)$ + h.c. survives and the O(a) corrections to this all cancel. The gauge coupling in (5.12) is an example of the type of irrelevant coupling one can have in lattice theory, i.e. one that goes to zero in the continuum limit. This may be seen by expanding (5.12) in powers of a : the O(a) terms vanish as already remarked, so the first correction is, in momentum space, $ga(\sin p_{\mu}a)\ [\bar{L}A_{\mu}\phi R]$. That is the current vanishes at zero external momentum and near the continuum. Noting this, the simpler interactions in (5.11) will be used henceforth rather than those in (5.12).

In summary, then, the spectrum is as follows:

(a) 3 massive gauge bosons and no residual gauge symmetry.

(b) Wilson fermion $\psi = L_2+R$, mass $\dfrac{1-8Kr}{2aK}$ – no doubling problem.

(c) Naive Weyl fermion L_1 which doubles to form an equal number of left and right massless states.

(d) Left-handed current coupling L_2 to itself and to L_1 with the gauge field $\underset{\sim}{A}_\mu$.

(e) Mixed (left and right) current coupling L_1 to itself and to $\underset{\sim}{A}_\mu$.

(f) Note that R, the right-handed part of ψ, has no gauge interactions. The only interactions it has are through the Higgs, and in the unitary gauge these have been gauged away to provide the mass terms for ψ.

(ii) SU(2) with fields in the adjoint representation

Both L and ϕ are real triplets $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$, $\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}$

and $(U_\mu(n))_{jk} = \exp\,[iga\, A^i_\mu\, \varepsilon_{ijk}]$, the antisymmetric objects ε_{ijk} being the generators of the adjoint representation of SU(2). Now $\dim G = 3$ as before but now

$$M_o = \{\phi : \phi_1^2 + \phi_2^2 + \phi_3^2 = 1\}$$

and so $\dim M_o = 2$. Hence (5.8) gives us $\dim H = 1$ and in fact $H = U(1)$. Therefore there will be a residual gauge symmetry with one of the gauge bosons remaining massless.

Let $\phi_o = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$: the gauge boson mass term is

$$- 2g^2\beta_H\, (\phi_o^\dagger)_j \left[A^i_\mu\, \varepsilon_{ijk}\, A^\ell_\mu\, \varepsilon_{\ell km}\right] (\phi_o)_m$$

$$= 2g^2\beta_H\, A^i_\mu\, A^\ell_\mu \left[\delta_{i\ell}\delta_{33} - \delta_{i3}\,\delta_{\ell 3}\right]$$

$$= \quad 2g^2\beta_H((A_\mu^1)^2 + (A_\mu^2)^2) \quad .$$

So A_μ^1 and A_μ^2 pick up the same mass each and A_μ^3 remains massless. The fermionic terms in the action are

$$\begin{aligned}
S_F = & - K \sum_{n,\mu} \left[\bar{L}_1(n)\gamma_\mu L_1(n+\mu) - \bar{L}_1(n+\mu)\gamma_\mu L_1(n)\right] \\
& - K \sum_{n,\mu} \left[\bar{L}_2(n)\gamma_\mu L_2(n+\mu) - \bar{L}_2(n+\mu)\gamma_\mu L_2(n)\right] \\
& - K \sum_{n,\mu} \left[\bar{L}_3(n)\gamma_\mu L_3(n+\mu) - \bar{L}_3(n+\mu)\gamma_\mu L_3(n)\right] \\
& - K \sum_{n,\mu} \left[\bar{R}(n)\gamma_\mu R(n+\mu) - \bar{R}(n+\mu)\gamma_\mu R(n)\right] \\
& - \sum_{n} \left[\bar{L}_3(n)R(n) + \bar{R}(n)L_3(n)\right] \\
& + Kr \sum_{n,\mu} \left[\bar{L}_3(n)R(n+\mu) + \bar{R}(n)L_3(n+\mu) + h.c.\right] \\
& - igK \sum_{n,\mu} \varepsilon_{ijk} A_\mu^i \left[\bar{L}_j(n)\gamma_\mu L_k(n+\mu) + \bar{L}_j(n+\mu)\gamma_\mu L_k(n)\right]
\end{aligned} \tag{5.13}$$

Putting $\psi = L_3 + R$ and setting the current

$$\underline{j}_\mu(n) = \underline{\bar{L}}(n)\gamma_\mu \times \underline{L}(n+\mu) + \underline{\bar{L}}(n+\mu)\gamma_\mu \times \underline{L}(n)$$

the action reduces to

$$S_F = -\sum_n \bar{\psi}(n)\psi(n) + K\sum_{n,\mu}\left[\bar{\psi}(n)(r-\gamma_\mu)\psi(n+\mu) + \bar{\psi}(n+\mu)(r+\gamma_\mu)\psi(n)\right]$$

$$- K\sum_{n,\mu}\left[\bar{L}(n)\gamma_\mu L_1(n+\mu) - \bar{L}_1(n+\mu)\gamma_\mu L_1(n)\right]$$

$$- K\sum_{n,\mu}\left[\bar{L}_2(n)\gamma_\mu L_2(n+\mu) - \bar{L}_2(n+\mu)\gamma_\mu L_2(n)\right]$$

$$- igK\sum_{n,\mu} \underset{\sim}{j}_\mu(n)\cdot\underset{\sim}{A}_\mu(n) \qquad (5.14)$$

and the spectrum may be summarized as follows:

(a) Two massive gauge bosons, each of the same mass and one massless one with the corresponding residual gauge symmetry.

(b) One Wilson fermion $\psi = L_3 + R$, mass $\frac{1-8Kr}{2aK}$.

(c) Two massless naive Weyl fermions (L_1, L_2) which will double, as before.

(d) A left-handed current coupling of L_3 to L_1 and L_2 through the massive gauge bosons A^2_μ and A^1_μ respectively.

(e) A mixed (left and right) current coupling L_1 to L_2 through the massless gauge boson A^3_μ - this is the "electromagnetic" current corresponding to the surviving gauge invariance.

Once again, $R = (\frac{1+\gamma_5}{2})\psi$ has no gauge interactions.

(iii) <u>SU(3) with fields in the fundamental representation</u>

L and ϕ are complex triplets $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$, $\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}$

and $U_\mu(n) = \exp[iga\, \underline{A}_\mu \cdot \underline{\lambda}]$ where the matrices $\{\lambda_a\}$ are the Gell-Mann generators of the fundamental representation of SU(3). This time $\dim G = 8$ and

$$M_o = \{\phi : |\phi_1|^2 + |\phi_2|^2 + |\phi_3|^2 = 1\}$$

so $\dim M_o = 5$. Hence from (5.8) $\dim H = 3$ and it turns out that H is SU(2). We should get three massless gauge bosons corresponding to the residual SU(3) and 5 massive ones.

Let $\phi_o = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$: the boson mass term is

$$2g^2\beta_H A_\mu^a A_\mu^b \,(0\ 0\ 1)\lambda^a \lambda^b \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= g^2\beta_H A_\mu^a A_\mu^b \,(0\ 0\ 1)\{\lambda^a,\lambda^b\} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

Now $\{\lambda^a, \lambda^b\} = \frac{4}{3} \delta^{ab}\, 1 + 2\, d^{abc} \lambda^c$

where d^{abc} is the symmetric invariant tensor of the group. The bottom right hand corner of λ^c is zero for all the Gell-Mann matrices except λ_8, for which this entry is $\frac{-2}{\sqrt{3}}$. Hence

$$(0\ 0\ 1)\{\lambda^a, \lambda^b\} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{4}{3}\delta^{ab} - \frac{4}{\sqrt{3}}\, d^{ab8} \tag{5.15}$$

d^{ab8} is diagonal with $d^{118} = d^{228} = d^{338} = \frac{1}{\sqrt{3}}$,

$d^{448} = d^{558} = d^{668} = d^{778} = -\frac{1}{2\sqrt{3}}$ and $d^{888} = -\frac{1}{\sqrt{3}}$.

Hence (5.15) becomes the diagonal matrix

$\mathrm{diag}(0,0,0,2,2,2,2,\frac{8}{3})$ and the mass term becomes

$$2g^2\beta_H \left[(A_\mu^4)^2 + (A_\mu^5)^2 + (A_\mu^6)^2 + (A_\mu^7)^2 + \frac{4}{3}(A_\mu^8)^2 \right] .$$

Hence A_μ^1, A_μ^2, A_μ^3 remain massless, with the surviving gauge invariance generated by the "isospin" SU(2) embedded in SU(3) and the remaining 5 bosons pick up masses as shown. The fermionic terms in the action are identical to those in (5.13) except for the last line, the current interaction, which is replaced by

$$-ig\ K \sum_{n,\mu} \left[\bar{L}(n)\gamma_\mu \lambda^a L(n+\mu) + \bar{L}(n+\mu)\gamma_\mu \lambda^a L(n) \right] \cdot A_\mu^a(n)$$

which is similar to the current interaction in model (i) (5.9). Putting $\psi = L_3 + R$ and the current

$$j_\mu^a(n) = \bar{L}(n)\gamma_\mu \lambda^a L(n+\mu) + \bar{L}(n+\mu)\gamma_\mu \lambda^a L(n)$$

we have formally the same action for the fermionic parts as (5.14). The spectrum is

(a) 5 massive gauge bosons and 3 massless ones, with the residual SU(2) gauge invariance.

(b) One Wilson fermion ψ, with mass $\dfrac{1 - 8Kr}{2aK}$.

(c) Two massless naive Weyl fermions as before.

(d) A current interaction $j_\mu^a\ A_\mu^a$ between L_1, L_2, L_3 and the bosons A_μ^a.

This current may be studied a little more deeply, to locate where the residual gauge invariance lies. If we put

$$A_\mu^\pm = A_\mu^1 \pm iA_\mu^2, \quad B_\mu^\pm = A_\mu^4 \pm iA_\mu^5, \quad C_\mu^\pm = A_\mu^6 \pm iA_\mu^7$$

then

$$\lambda^a A^a_\mu = \begin{pmatrix} A^3_\mu & A^-_\mu & B^-_\mu \\ A^+_\mu & -A^3_\mu & C^-_\mu \\ B^+_\mu & C^+_\mu & \frac{-2}{\sqrt{3}} A^8_\mu \end{pmatrix} .$$

So L_1 and L_2 couple to themselves and each other with A^3_μ and $A^\pm_\mu$ respectively: this is where the residual SU(2) gauge invariance lies. This will not be a left-handed current as L_1, L_2 lose their left-handed nature. L_3 couples to itself with A^8_μ and to L_1 and L_2 with $B^\pm_\mu$ and $C^\pm_\mu$ respectively: these are purely left-handed as L_3 retains its left-handedness

As before, R has no gauge interactions.

These three models all have a common feature: the existence of one or more naive Weyl fermions which undergo doubling and so lose their left-handedness: thus the current interaction is not genuinely left-handed for all the fields. This problem can be surmounted at once and we give next an example of how this is done in model (i).

(iv) <u>SU(2), fundamental representation and two right-handed fields</u>

We introduce another right handed singlet R_1 with kinetic term

$$- K \sum_{n,\mu} \left[\bar{R}_1(n)\gamma_\mu R_1(n+\mu) - \bar{R}_1(n+\mu)\gamma_\mu R_1(n) \right]$$

and add this to the action (5.6) along with the interaction terms

$$Kr' \sum_{n,\mu} \left[\bar{L}(n)\tilde{\phi}(n)R_1(n+\mu) + \bar{R}_1(n)\tilde{\phi}^\dagger(n+\mu)L(n+\mu) + \text{h.c.} \right]$$

$$- \sum_n \left[\bar{L}(n)\tilde{\phi}(n)R_1(n) - \bar{R}_1(n)\tilde{\phi}^\dagger(n)L(n) \right] \qquad (5.16)$$

where $\tilde{\phi}(n)$ is the charge conjugate field $(\tilde{\phi}(n) = i\tau_2(\phi^\dagger(n))^t$ for SU(2)). The interactions (5.16) are the analogues of (5.5). Upon introducing the symmetry breaking $\phi_o = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ as in model (i), here we have $\tilde{\phi}_o = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ as well and so (5.9) picks up the extra terms from (5.16) as follows

$$Kr' \sum_{n,\mu} \Big[\bar{L}_1(n)R_1(n+\mu) + \bar{R}_1(n)L_1(n+\mu) + \bar{L}_1(n+\mu)R_1(n) + \bar{R}_1(n+\mu)L_1(n) \Big]$$

$$- \sum_n \Big[\bar{L}_1(n)R_1(n) + \bar{R}_1(n)L_1(n) \Big]$$

along with the kinetic term for R_1 above. Collecting these together and putting $\chi = L_1 + R_1$ (5.10) becomes

$$\begin{aligned} S_F = & - \sum_n \bar{\psi}(n)\psi(n) + K \sum_{n,\mu} \Big[\bar{\psi}(n)(r-\gamma_\mu)\psi(n+\mu) + \bar{\psi}(n+\mu)(r+\gamma_\mu)\psi(n) \Big] \\ & - \sum_n \bar{\chi}(n)\chi(n) + K \sum_{n,\mu} \Big[\bar{\chi}(n)(r'-\gamma_\mu)\chi(n+\mu) + \bar{\chi}(n+\mu)(r'+\gamma_\mu)\chi(n) \Big] \\ & - igK \sum_{n,\mu} \underline{j}_\mu(n) \cdot \underline{A}_\mu(n) \end{aligned} \qquad (5.17)$$

where $\underline{j}_\mu(n) = \bar{L}(n)\gamma_\mu \underline{\tau} L(n+\mu) + \bar{L}(n+\mu)\gamma_\mu \underline{\tau} L(n)$

the same current as before. The spectrum of this model is as follows:-

(a) As the group theory is the same as model (i), there are three massive bosons, degenerate in mass.

(b) There are two Wilson fermions, ψ and χ, with masses $\frac{1 - 8Kr}{2aK}$ and $\frac{1 - 8Kr'}{2aK}$ respectively.

(c) A current $\underline{j}_\mu$ which is genuinely left handed (as it is composed of $L_1 = (\frac{1-\gamma_5}{2})\psi$ and $L_2 = (\frac{1-\gamma_5}{2})\chi$) which couples to the bosons $\underline{A}_\mu$.

(d) There is no right handed current at all: $\underline{j}_\mu(n)$ is left handed only.

These four models have two features in common. The right handed parts of the Wilson fermions have no gauge interactions and indeed will not be able to pick any up in renormalization; and the second feature is that we have had to use spontaneous symmetry breaking so there are massive gauge bosons. Now our motivation was to put neutrinos onto the lattice by ensuring that only their left handed parts interacted with the gauge field, so the first feature is the one we are searching for. However, the second feature means that we have to do this within the context of spontaneous symmetry breaking so we are led directly on to consider the Glashow-Salam-Weinberg theory; the massive bosons will then have to be the W and Z of the standard electroweak theory.

5.5 The Electroweak Theory on the Lattice

Using the ideas of the previous section we can now formulate a model of the $SU(2) \times U(1)$ gauge theory on the lattice. We now consider first just the leptonic first generation sector: electrons and electron neutrinos. We begin by defining the fields, their quantum numbers and their gauge transformation properties. (For a review of the continuum model, see J.C. Taylor (1978) and H. Fritsch and P. Minkowski (1981)). L, ϕ and $\tilde{\phi}$ are complex doublets under SU(2) and the usual singlets under U(1) with various values of the associated hypercharge. R_1 transforms trivially under SU(2) (zero weak isospin) but has the usual U(1) property and R_2 transforms trivially under both SU(2) and U(1) (zero weak isospin and zero hypercharge). See Table 5.1.

$\Omega(n)$ is a member of the fundamental representation of SU(2) and represents SU(2) gauge transformations, $e^{i\alpha(n)}$ is a member of some representation of U(1) and represents U(1) gauge transformations. The gauge transformation of the conjugate fields $\overline{L}$, $\phi^\dagger$ etc. follow from those in Table 5.1. Note that $\underset{\sim}{\phi}$ is not a dynamically independent object:

$$\tilde{\phi} \quad = \quad i\,\tau_2(\phi^\dagger)^t \quad = \quad i\,\tau_2\phi^* \quad = \quad \begin{pmatrix} \phi_o^\dagger \\ -\phi_+^\dagger \end{pmatrix}$$

and so the gauge transformation properties can be found from that of ϕ.

$$\begin{aligned} \phi(n) \quad &= \quad i\,\underline{\tau}_2(\phi^\dagger(n))^t \\ &\rightarrow i\,\tau_2(\phi^\dagger(n)\;e^{-\frac{1}{2}i\alpha(n)}\;\Omega^\dagger(n))^t \\ &= \quad i\,\underline{\tau}_2(\Omega^\dagger(n))^t\;e^{-\frac{1}{2}i\alpha(n)}\;(\phi^\dagger(n))^t \end{aligned} \tag{5.18}$$

Now $\Omega(n)$ is a 2×2 special unitary matrix: hence

$$\tau_2\Omega^* \quad = \quad \Omega\tau_2$$

as may be verified for example by putting

$$\Omega \quad \doteq \quad \begin{pmatrix} \cos\theta\;e^{i\phi} & \sin\theta\;e^{i\psi} \\ -\sin\theta\;e^{-i\psi} & \cos\theta\;e^{-i\phi} \end{pmatrix}$$

which is the general form for $\Omega \in SU(2)$.

Consequently, (5.18) becomes

Fields	SU(2) Weak Isospin I	SU(2) Weak Isospin I_3	U(1) Hypercharge Y	Electromagnetic Charge Q $= I_3 + Y$	Properties under gauge transformation $\Omega(n) \varepsilon SU(2)$, $e^{i\alpha(n)} \varepsilon U(1)$
$L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	$\frac{1}{2}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$L(n) \rightarrow \Omega(n) e^{-\frac{1}{2} i\alpha(n)} L(n)$
$R_1 = e_R$	0	0	-1	-1	$R_1(n) \rightarrow e^{-i\alpha(n)} R_1(n)$
$R_2 = \nu_R$	0	0	0	0	$R_2(n) \rightarrow R_2(n)$
$\phi = \begin{pmatrix} \phi_+ \\ \phi_o \end{pmatrix}$	$\frac{1}{2}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$+\frac{1}{2}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\phi(n) \rightarrow \Omega(n) e^{+\frac{1}{2} i\alpha(n)} \phi(n)$
$\tilde{\phi} = \begin{pmatrix} \phi_o^\dagger \\ -\phi_+^\dagger \end{pmatrix}$	$\frac{1}{2}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\tilde{\phi}(n) \rightarrow \Omega(n) e^{-\frac{1}{2} i\alpha(n)} \tilde{\phi}(n)$

TABLE 5.1

Matter Fields: quantum numbers and gauge transformation properties.

$$\tilde{\phi}(n) \rightarrow i\ \Omega(n)\ \tau_2\ e^{-\frac{1}{2}i\alpha(n)}(\phi^\dagger(n))^t$$

$$= e^{-\frac{1}{2}i\alpha(n)}\ \Omega(n)\ \tilde{\phi}(n)$$

as required (see Table 5.1).

The gauge fields are given in Table 5.2. The SU(2) fields are designated $U_\mu(n)$ and are straightforward. The U(1) fields are designated $V_\mu(n)$, but since we are using different representations of U(1), we need more than one gauge representation: $V'(n)$ is used to accommodate fields with a hypercharge of 1 rather than $\frac{1}{2}$. The gauge transformation properties are given in Table 5.2. It should be stressed that V_μ and V'_μ are not dynamically independent objects; they are representations of the <u>same</u> fundamental object and are expressed in terms of just one independent object on each link:- $B_\mu(n)$. In fact, $V'_\mu(n) = [V_\mu(n)]^2$.

We are now ready to construct the action. The fermionic kinetic parts are

$$\begin{aligned} S_K = & -K \sum_{n,\mu}\left[\overline{L}(n)\gamma_\mu U_\mu(n)V_\mu(n)L(n+\mu) - \overline{L}(n+\mu)\gamma_\mu U_\mu^\dagger(n)V_\mu^\dagger(n)L(n)\right] \\ & -K \sum_{n,\mu}\left[\overline{R}_1(n)\gamma_\mu V'_\mu(n)R_1(n+\mu) - \overline{R}_1(n+\mu)\gamma_\mu V'^\dagger_\mu(n)R_1(n)\right] \\ & -K \sum_{n,\mu}\left[\overline{R}_2(n)\gamma_\mu R_2(n+\mu) - \overline{R}_2(n+\mu)\gamma_\mu R_2(n)\right] \end{aligned} \tag{5.19}$$

This is clearly gauge invariant, as can be checked by using the gauge transformations listed in Tables 5.1 and 5.2. The interaction terms, which will give the fermions their masses after spontaneous symmetry breaking are

Fields	I	Y	Gauge transformation
$U_\mu(n) = \exp[\frac{1}{2} iga\, \underline{\tau} \cdot \underline{A}_\mu(n)]$	$\frac{1}{2}$	0	$U_\mu(n) \to \Omega(n) U_\mu(n) \Omega^{-1}(n+\mu)$
$V_\mu(n) = \exp[-\frac{1}{2}\, ig'aB_\mu(n)]$	0	$-\frac{1}{2}$	$V_\mu(n) \to e^{-\frac{1}{2}i\alpha(n)}\, V_\mu(n)\, e^{\frac{1}{2}i\alpha(n+\mu)}$
$V'_\mu(n) = \exp[-ig'a\, B_\mu(n)]$	0	-1	$V_\mu{}'(n) \to e^{-i\alpha(n)}\, V'_\mu(n)\, e^{-i\alpha(n+\mu)}$

TABLE 5.2

Gauge fields: quantum numbers and gauge transformation properties.

$$S_I = Kr \sum_{n,\mu} \left[\bar{L}(n)\phi(n)V'_\mu(n)R_1(n+\mu) + \bar{R}_1(n)V'_\mu(n)\phi^\dagger(n+\mu)L(n+\mu) + \text{h.c.} \right]$$

$$+ Kr' \sum_{n,\mu} \left[\bar{L}(n)\tilde{\phi}(n)R_2(n+\mu) + \bar{R}_2(n)\tilde{\phi}^\dagger(n+\mu)L(n+\mu) + \text{h.c.} \right]$$

$$- \sum_{n} \left[\bar{L}(n)\phi(n)R_1(n) + \bar{R}_1(n)\phi^\dagger(n)L(n) \right]$$

$$- \sum_{n} \left[\bar{L}(n)\tilde{\phi}(n)R_2(n) + \bar{R}_2(n)\tilde{\phi}^\dagger(n)L(n) \right] \tag{5.20}$$

Finally there are the kinetic parts for ϕ and for the gauge fields. Since $\tilde{\phi}$ is not dynamically independent, as discussed earlier, it is covered by the kinetic term for ϕ.

$$S_H = \beta_H \sum_{n,\mu} \left[\phi^\dagger(n)U_\mu(n)V^\dagger_\mu(n)\phi(n+\mu) + \phi^\dagger(n+\mu)U^\dagger_\mu(n)V_\mu(n)\phi(n) \right] \tag{5.21}$$

$$S_G = \beta \,\text{Re Tr} \sum_n \left[\prod_{\text{plaq}} U_\mu \right] + \beta' \,\text{Re Tr} \sum_n \left[\prod_{\text{plaq}} V_\mu \right] . \tag{5.22}$$

Again, it is straightforward to check the gauge invariance of equations (5.20) to (5.22). Note that since ϕ has hypercharge $+\frac{1}{2}$ rather than $-\frac{1}{2}$, we have to use $V^\dagger_\mu$ in the place of V_μ and vice versa. The full action is then

$$S = S_K + S_I + S_H + S_G .$$

We now examine the spectrum in the Higgs phase in the same way as in the last section. Set $\phi_o = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ so $\tilde{\phi}_o = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and put $e = e_L + e_R$ and $\nu = \nu_L + \nu_R$. Then the fermions pick up momentum dependent masses from (5.20), and (5.19) and (5.20) together give

$$K \sum_{n,\mu} \left[\bar{e}(n)(r-\gamma_\mu)e(n+\mu) - \bar{e}(n+\mu)(r+\gamma_\mu)e(n) \right] - \sum_n \bar{e}(n)e(n)$$

$$+ K \sum_{n,\mu} \left[\bar{\nu}(n)(r'-\gamma_\mu)\nu(n+\mu) - \bar{\nu}(n+\mu)(r'+\gamma_\mu)\nu(n) \right] - \sum_n \bar{\nu}(n)\nu(n) \qquad (5.23)$$

as the fermionic kinetic terms. We have, then, two Wilson fermions with masses

$$m_e = \frac{1 - 8Kr}{2aK}, \qquad m_\nu = \frac{1 - 8Kr'}{2aK}. \qquad (5.24)$$

These can be tuned to their physical values as we take the continuum limit - in particular we would want $K \to \frac{1}{8r'}$ (for the free theory) if we require the neutrino to be massless.

The boson mass term comes from (5.21) and is identical to its continuum analogue: we get

$$\frac{\beta_H a^2}{4} \left[g^2 ({A^1}^2 + {A^2}^2) + g^2 {A^3_\mu}^2 + g'^2 B^2_\mu - 2gg' A^3_\mu B_\mu \right]. \qquad (5.25)$$

Putting $W_\mu = \dfrac{A^1_\mu - iA^2_\mu}{\sqrt{2}}$, $g'/g = \tan\theta$ and

$$\begin{aligned} A_\mu &= A^3_\mu \sin\theta + B_\mu \cos\theta \\ Z_\mu &= A^3_\mu \cos\theta - B_\mu \sin\theta \end{aligned} \qquad (5.26)$$

(5.25) becomes

$$a^2\beta_H(\tfrac{1}{2}g^2 W_\mu \bar{W}_\mu + \frac{g^2}{4\cos^2\theta} Z_\mu Z_\mu)$$

$$\equiv m_W^2 \bar{W}_\mu W_\mu + \tfrac{1}{2}m_Z^2 Z_\mu Z_\mu .$$

This is the familiar structure of the standard theory: the W boson is charged, the Z neutral and θ is the weak mixing angle.

The photon remains massless, being the gauge particle of the surviving U(1) invariance.

There remains the current interaction arising from the gauge coupling in (5.19). This may be written

$$- iK \sum_{n,\mu} \left[g \; \underline{j}_\mu(n) \cdot \underline{A}_\mu(n) - g' \; j_\mu^Y(n) \; B_\mu(n) \right]$$

where $\underline{j}_\mu(n)$ is the weak isospin current

$$\underline{j}_\mu(n) \quad = \quad \tfrac{1}{2}(\overline{L}(n)\gamma_\mu \; \underline{\tau} \; L(n+\mu) \; + \; \overline{L}(n+\mu)\gamma_\mu \; \underline{\tau} \; L(n))$$

and $j_\mu^Y(n)$ is the hypercharge current

$$j_\mu^Y(n) \quad = \quad \tfrac{1}{2}(\overline{L}(n)\gamma_\mu L(n+\mu) \; + \; \overline{L}(n+\mu)\gamma_\mu L(n))$$

$$+ \; (\overline{R}_1(n)\gamma_\mu R_1(n+\mu) \; + \; \overline{R}_1(n+\mu)\gamma_\mu R_1(n)) \quad .$$

Once again, this structure is identical to that of the continuum model so the currents derived will be precisely what is required. Putting

$$J_\mu^\ell \quad = \quad 2(j_\mu^1 + ij_\mu^2) \quad = \quad \overline{\nu}\gamma_\mu(1-\gamma_5)e$$

we have

$$j_\mu^1 \; A_\mu^1 \; + j_\mu^2 \; A_\mu^2 \quad = \quad \frac{1}{2\sqrt{2}} \; (J_\mu^\ell \; W_\mu \; + \; J_\mu^{\ell\dagger} \; \overline{W_\mu})$$

and so the current coupling to the W is

$$- \frac{iKg}{\sqrt{2}} \sum_{n,\mu} \left[\overline{\nu}(n)\gamma_\mu(\frac{1 - \gamma_5}{2}) \; e(n+\mu) \cdot W_\mu(n) \right.$$

$$\left. + \; \overline{e}(n)\gamma_\mu(\frac{1 - \gamma_5}{2})\nu(n+\mu) \cdot \overline{W_\mu}(n) \right] \; + \; \text{h.c.} \qquad (5.28)$$

So J_μ^ℓ is the charged leptonic weak current, coupling to the W boson just as it should. The remaining terms in (5.27) couple to A_μ^3 and B_μ, so with the aid of (5.26) we can derive the coupling to the photon and to the Z boson. To the photon we have

$$g \sin\theta \; j_\mu^3 - g' \cos\theta \; j_\mu^Y = - g \sin\theta \; \bar{e} \gamma_\mu e$$

giving us the current coupling to the photon

$$iKg \sin\theta \sum_{n,\mu} \left[\bar{e}(n)\gamma_\mu \; e(n+\mu) \cdot A_\mu(n) \right] + \text{h.c.} \tag{5.29}$$

(5.29) is just the electromagnetic coupling and so does not - and indeed should not - involve the neutrino which has zero electromagnetic charge. This current has no left-right asymmetry.

The coupling to the Z is through the current

$$g \cos\theta \; j_\mu^3 + g' \sin\theta \; j_\mu^Y$$

$$= \bar{\nu}_L \gamma_\mu \nu_L \left(\frac{1+\tan^2\theta}{2}\right) - \bar{e}_L \gamma_\mu e_L \left(\frac{1-\tan^2\theta}{2}\right) + \bar{e}_R \gamma_\mu e_R \cdot \tan^2\theta$$

giving us the current coupling

$$-iKg \cos\theta \sum_{n,\mu} \Bigg\{ \left(\frac{1+\tan^2\theta}{2}\right) \bar{\nu}(n)\gamma_\mu \left(\frac{1-\gamma_5}{2}\right) \nu(n+\mu) \cdot Z_\mu(n)$$

$$+ \bar{e}(n)\gamma_\mu \left[\tan^2\theta \left(\frac{1+\gamma_5}{2}\right) - \frac{1-\tan^2\theta}{2} \left(\frac{1-\gamma_5}{2}\right) \right] e(n+\mu) \cdot Z_\mu(n) \Bigg\} + \text{h.c.} \tag{5.30}$$

This is the same as the continuum current - in particular we see that the Z couples to the neutrino only through a left-handed current and that the Z couples to the electron through a left-right asymmetric current, just as in the continuum model.

This completes the analysis of the spectrum. In summary, we have the electron and neutrino, both Wilson fermions, with definite masses given by (5.24). The gauge bosons acquire masses and since the group theory is the same as in the continuum model this also happens in the same way as in the continuum.

The current couplings to the bosons are also the same as in the continuum theory. The neutrino only couples through left-handed currents: the right handed part has no gauge interactions whatsoever. Naturally enough, the naive continuum limit of this model is precisely the Weinberg-Salam model for electrons and neutrinos.

5.6 A Low Energy Theory

We may derive a low energy theory from the Higgs regime of the action presented in the last section in much the same way that a low energy theory - the Fermi theory - may be derived from the continuum electroweak model. We do a large β_H expansion and integrate out the W and Z fields (see also Bowler et al. (1981)) (Since the boson masses are proportional to β_H, an expansion in β_H large is equivalent to a large mass expansion.) Consider first the W_μ fields. Writing

$$J_\mu(n) = 2iK\left[\bar{\nu}(n)\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)e(n+\mu) + \bar{\nu}(n+\mu)\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)e(n)\right]$$

the parts of the action involving W_μ may be expanded to give

$$e^{-m_W^2 \bar{W}_\mu W_\mu} \left[1 + \frac{g}{2\sqrt{2}} (J_\mu W_\mu + J_\mu^\dagger \overline{W_\mu}) + \frac{g^2}{8} (J_\mu J_\mu^\dagger W_\mu \overline{W_\mu}) \right.$$

$$\left. + \frac{g^2}{16} (J_\mu^2 W_\mu^2 + J_\mu^{\dagger 2} W_\mu^2) + \ldots\ldots \right] \qquad (5.31)$$

where we have expanded only to quadratic terms.

Defining $\langle A \rangle = \int \left[\prod_{n,\mu} \prod_{m,\nu} dW_\mu(n)\, d\bar{W}_\nu(m) \right] e^{-\sum_{n,\mu} m_W^2 \bar{W}_\mu(n) W_\mu(n)} \cdot A$

we shall use the following results:-

$$\langle 1 \rangle = C, \qquad \langle W_\mu \rangle = \langle \bar{W}_\mu \rangle = 0,$$

$$\langle W_\mu^2 \rangle = \langle \bar{W}_\mu^2 \rangle = 0,$$

$$\langle W_\mu \bar{W}_\mu \rangle = C \frac{1}{m^2{}_W}$$

where C is some (infinite) constant.

We integrate out the W fields in (5.31) to obtain

$$C \left[1 + \frac{g^2}{8m_W^2} (J_\mu J_\mu^\dagger) + 0 \left(\frac{1}{m_W^4}\right) \right] \qquad (5.32)$$

which may be re-exponentiated to contribute towards the effective action. Now let us turn our attention to the Z fields. Putting

$$J_\mu^N(n) = 2iK\cos^2\theta \left[\left(\frac{1+\tan^2\theta}{2}\right) \cdot \left[\bar{\nu}(n)\gamma_\mu \left(\frac{1-\gamma_5}{2}\right)\nu(n+\mu) + \bar{\nu}(n+\mu)\gamma_\mu \left(\frac{1-\gamma_5}{2}\right)\nu(n) \right] \right.$$

$$- \left(\frac{1-\tan^2\theta}{2}\right) \left[\bar{e}(n)\gamma_\mu \left(\frac{1-\gamma_5}{2}\right) e(n+\mu) + \bar{e}(n+\mu)\gamma_\mu \left(\frac{1-\gamma_5}{2}\right) e(n) \right]$$

$$\left. + \tan^2\theta \left[\bar{e}(n)\gamma_\mu \left(\frac{1+\gamma_5}{2}\right) e(n+\mu) + \bar{e}(n+\mu)\gamma_\mu \left(\frac{1+\gamma_5}{2}\right) e(n) \right] \right]$$

the action involving Z_μ may be expanded (again only to second order) to give

$$e^{-\frac{1}{2}m_Z^2 Z^2}\left[1 + \frac{g}{2\cos\theta} J_\mu^{\,N} Z_\mu + \frac{g^2}{8\cos^2\theta} J_\mu^{\,N} J_\mu^{\,N} Z_\mu Z_\mu + D\cdot Z_\mu^2 + \ldots\right]. \quad 5.33)$$

The last term in (5.33) requires some explanation. We are expanding the exponential e^{-S} to second order in the gauge fields - but there are some terms in S which are already second order. These come from second order expansions of the gauge variables U_μ and V_μ. All such expansions in the case of the W fields give rise to irrelevant currents of the type described earlier, and so were not included in (5.31) and (5.32). However, for the Z field there is one term which produces a current not of the irrelevant type. It comes from the expansion to second order of V'_μ in the first line of (5.20), which gives, after setting $\phi_o = \begin{pmatrix}0\\1\end{pmatrix}$,

$$-\frac{Kr\, g'^2}{2} \sum_{n,\mu} [\bar{e}(n)e(n+\mu) + \bar{e}(n+\mu)e(n)]\ B_\mu^2(n)\,.$$

This may be expanded as the other terms are - but clearly it gives quadratic terms in the boson field after expansion of e^{-S} only to the linear term. Thus we may rewrite D, using (5.26)

$$D(n) = \tfrac{1}{2}Kr\ g^2\tan^2\theta\sin^2\theta\left[\bar{e}(n)e(n+\mu) + \bar{e}(n+\mu)e(n)\right] .$$

As before, define

$$\langle A\rangle = \int \left[\prod_{n,\mu} dZ_\mu(n)\right] e^{-\sum_{n,\mu} \frac{1}{2}m_Z^2 Z_\mu^2(n)} \cdot A \ .$$

The analogue results are

$$<1> \quad = \quad C', \qquad\qquad <Z_\mu> \quad = \quad 0 ,$$

$$<Z_\mu^2> \quad = \quad C' \frac{1}{m_Z^2}$$

(C' is another infinite constant).

Integrating out the Z fields in (5.33) we obtain

$$C' \left[1 + \frac{g}{8m_Z^2\cos^2\theta} (J_\mu{}^N J_\mu{}^N) + \frac{D}{m_Z^2} + O(\frac{1}{m_Z^4}) \right] \qquad (5.34)$$

This, too, may be re-exponentiated, and when taken together with (5.32) gives us the effective interaction action S_I^{eff} (note $m_Z \cos\theta = m_W$)

$$S_I^{eff} = \frac{g^2}{8m_W^2} \left[J_\mu J_\mu^\dagger + J_\mu{}^N J_\mu{}^N \right] + O(\frac{1}{m_W^4}) \quad . \qquad (5.35)$$

D does not participate in the interaction action, but must be absorbed into the electron kinetic term to provide a renormalization of the electron mass parameter r. The renormalizing term is

$$\frac{D}{m_Z^2} = \frac{D\cos^2\theta}{m_W^2}$$

$$= \frac{1}{2} \frac{g^2}{m_W^2} \sin^4\theta \cdot Kr \left[\bar{e}(n)e(n+\mu) + \bar{e}(n+\mu)e(n) \right] \quad .$$

and so the renormalisation is

$$r \rightarrow r\left[1 - \frac{g^2}{2m_W^2} \sin^4\theta + O(\frac{1}{m_W^4}) \right] .$$

Note that there is no renormalization of the neutrino mass - there are no corresponding self energy diagrams which can produce a $\bar{\nu}_L \nu_R$ type term, since the ν_R participates in no gauge interactions.

(5.35) is clearly the lattice analogue of the Fermi interactions, as indeed the naive continuum limit of this is indeed the Fermi theory.

5.7 Conclusions

All the left-right asymmetric models in this chapter share one general feature. The Higgs mechanism is invoked to generate momentum dependent mass terms which lift the degeneracy of the doubled states by giving the states at the edge of the Brillouin Zone a mass of the order of the cutoff. Thus the explicitly left handed currents associated with the low-lying state remain left handed and the chiral physics is preserved. However, it is not clear whether or not it is a coincidence of the construction that the existence of the left-right asymmetry has, of necessity, to be within a Higgs phase. We can do no more than speculate whether or not this is pointing to a deep connection between chiral physics and the Higgs mechanism.

We have presented a candidate model for the Weinberg-Salam model on the lattice. It possesses all the features of the continuum model, including the corresponding low energy Fermi theory. The right handed part of the neutrino, which we had to introduce, possesses no gauge interactions and indeed only interacts through a mass term with the left handed part.

The masses of the fermions (5.24) need to be tuned to their physical values, but this may be no less aesthetically pleasing than the continuum model in which the right handed part of the neutrino is omitted by fiat.

The other leptons and the quarks can be incorporated in the obvious way to provide a model of the $SU(3) \times SU(2) \times U(1)$ standard theory. It is only necessary to provide extra right-handed parts for the neutrinos, as the quarks already have them. The right handed parts of the quarks will have non-zero hypercharge (unlike the neutrinos) and so will participate in interactions with the photon and the Z boson, just as they do in the continuum model.

The continuum limit of the model in section 5.5 contains an anomaly. However, the lattice model must have anomaly cancelling; consequently one might expect the high mass (doubled) states to bring this about by their current interactions. Unfortunately, this may mean that the low energy physics (by which we mean the physics of the lowest lying states) may be influenced by these high mass states.

However, the full model, containing quarks and leptons, of the $SU(3) \times SU(2) \times U(1)$ gauge group has no anomalies in the continuum as they all cancel. Consequently, in the lattice model, one might expect the high mass states' influence on low energy physics to cancel in the same way. Since the continuum model is anomaly free one would hope that no extra problems could appear on the lattice.

Finally, there is now no obstruction, in principle, to constructing lattice models utilizing the SU(5) and high unification groups. All such models need a left-right asymmetric construction and it is possible that this may now be achieved.

REFERENCES

Atiyah, M. and Ward, R. (1977) Comm. Math. Phys. 55, 117.

Becher, P. and Joos, H. (1982) Z. Phys. C15, 343.

Belavin, A., Polyakov, A., Schwartz, A. and Tyupkin, Y., (1975) Phys. Lett. 59B, 85.

Binder, K., editor (1979), "Monte Carlo Methods in Statistical Physics" (Springer, New York).

Bogomol'nyi, E. and Fateev, V. (1977) Phys. Lett. 71B, 93.

Bott, R. (1956) Bull. Soc. Math. France 84, 251.

Bowler, K., Pawley, G., Pendleton, B. and Wallace, D. (1981) Phys. Lett. 104B, 481.

Bruce, A. and Wallace, D. (1981) Phys. Rev. Lett. 47, 1743.

Callan, C. (1980) Phys. Rep. 67, 123.

Callan, C., Dashen, R. and Gross, D. (1976) Phys. Lett. 63B, 334.

Callan, C., Dashen, R. and Gross, D. (1978) Phys. Rev. D17, 2717.

Callan, C., Dashen, R. and Gross, D. (1979) Phys. Rev. D20, 3279.

Coleman, S. (1977) Lectures given at the 1977 Erice Summer School, edited by A. Zichichi (Plenum, New York).

Creutz, M. (1980) Phys. Rev. D21, 1006.

Dyson, F. (1952) Phys. Rev. 85, 631.

Feynman, R. (1948) Rev. Mod. Phys. 20, 367.

Fradkin, E. (1959) Dokl. Akad. Sci. USSR 125, 321.

Fradkin, E. and Shenker, S. (1979) Phys. Rev. D19, 3682.

Fritsch, H. and Minkowski, P. (1981) Phys. Rep. 73, 67.

Griffiths, R. (1972) in "Phase Transitions and Critical Phenomena" edited by C. Domb and M. Green (Academic, New York, Volume 3).

Ishikawa, K., Schierholz, G., Schneider, H. and Teper, M. (1983) Phys. Lett. 128B, 309.

REFERENCES (Contd.)

Jackiw, R., Nohl, C. and Rebbi, C. (1977) Lectures given at the 1977 Banff School (Plenum, New York).

Jackiw, R. and Rebbi, C. (1976) Phys. Rev. Lett. 37, 172.

Jaffe, A. (1965) Comm. Maths. Phys. 1, 127.

Jongeward, G., Stack, J. and Jayakaprash, C. (1980) Phys. Rev. D21, 3360.

José, J., Kadanoff, L., Kirkpatrick, S. and Nelson, D. (1977) Phys. Rev. B16, 1217.

Kadanoff, L. (1976) in "Phase Transitions and Critical Phenomena" edited by C. Domb and M. Green (Academic, New York, Volume 5a).

Karsten, L. and Smit, J. (1981) Nucl. Phys. B183, 103.

Kawamoto, N. and Smit, J. (1981) Nucl. Phys. B192, 100.

Kogut, J. (1979) Rev. Mod. Phys. 51, 659.

Kogut, J. (1983) Rev. Mod. Phys. 55, 775.

Kogut, J. and Susskind, L. (1975) Phys. Rev. D11, 395.

Lichtenberg, D. (1978) "Unitary Symmetry in Particle Physics" (2nd Edition, Academic, New York).

McLaughlin, D. (1972) J. Math. Phys. 13, 1099.

Nakomo, T. (1959) Prog. Theor. Phys. 21, 241.

Nielsen, H. and Ninomiya, M. (1981)(a) Nucl. Phys. B185, 20.

Nielsen, H. and Ninomiya, M. (1981)(b) Nucl. Phys. B193, 173.

Rabin, J. (1982) Nucl. Phys. B201, 315.

Schwinger, J. (1959) Phys. Rev. 115, 721.

Shore, G. (1979) Ann. Phys. 122, 321.

Taylor, J.C. (1978) "Gauge Theories of Weak Interactions" (Cambridge University Press.

't Hooft, G. (1976) Phys. Rev. D14, 3432.

Wilson, K. (1974) Phys. Rev. D10, 2455 also in "New Phenomena in Subnuclear Physics" edited by A. Zichchi, Erice, 1975 (Plenum, New York).